

Patent No. 7,940,652  
Petition for *Inter Partes* Review

**UNITED STATES PATENT AND TRADEMARK OFFICE**

---

**BEFORE THE PATENT TRIAL AND APPEAL BOARD**

---

Juniper Networks, Inc.  
Petitioner

v.

Brixham Solutions, LTD.  
Patent Owner

---

Case IPR: Unassigned

---

---

**DECLARATION OF TAL LAVIAN, Ph.D.**

I, Tal Lavian, declare as follows:

1. I have personal knowledge of the facts stated in this declaration, and could and would testify to these facts under oath if called upon to do so.

**I. INTRODUCTION**

2. I have been retained by counsel for Juniper Networks, Inc. (“Juniper”) in this case as an expert in the relevant art.

3. I have been asked to provide my opinions on the question of validity of **claims 1-5, 8-11, 13-15, and 17** (“Challenged Claims”) of **U.S. Patent No. 7,940,652** by Ping Pan (“the ’652 patent”), which is owned by Brixham Solutions, LTD (“Patent Owner” or “BSL”). The opinions discussed below are my own. In formulating these opinions, I have reviewed a variety of materials and made use of my own personal knowledge. The materials I have relied on in formulating my opinions are identified in this report, including in the attached Appendix List.

4. I am being paid \$400 per hour in connection with my work in this case. My compensation is not contingent on my reaching any particular findings or conclusions, or any outcome of the case.

## II. EXECUTIVE SUMMARY

1. On a high level, the '652 patent describes a technique for protecting data traffic on a Pseudowire that involves configuring a standby Pseudowire by sending a configuration parameter between a source node and a destination node that includes a “priority” for the standby Pseudowire, and then using that priority, at least in part, to make a determination as to whether to preempt existing traffic on the standby Pseudowire.

2. As described in more detail below, the concepts of configuring a standby Pseudowire, using configuration parameters to assign a priority to a Pseudowire, and preempting existing traffic based on priorities had been known in the art for many years prior to the '652 patent, and had been described in numerous industry standard documents, text books, and patents.

3. More specifically, it is my opinion that claims 1-5, 8-11, 13-15, and 17 of the '652 patent are anticipated by many prior art references, including **U.S. Patent Pub. No. 2004/0156313** to Hofmeister et al., **Request for Comments 3386** and/or “**Metro Ethernet**” by Sam Halabi.

4. If certain aspects recited in claims 1-5, 8-11, 13-15, and 17 of the '652 patent are not deemed to be disclosed or inherent over these references, then claims 1-5, 8-11, 13-15, and 17 of the '652 patent are certainly obvious in view of some combination of these references and/or in combination with **U.S. Patent No. 7,804,767 B1** to Owens et al., **Request for Comments 3209**, “**The LSP Protection/Restoration Mechanism in GMPLS**” by Ziyang Chen, **U.S. Patent No. 7,305,481 B2** to Blanchet et al. and/or **U.S. Patent Pub. No. 2006/0047851 A1** to Voit et al.

5. The bases for my opinions are set forth in detail below.

### **III. BACKGROUND AND QUALIFICATIONS**

7. I possess the knowledge, skills, experience, training and the education to form an expert opinion and testimony in this case. A detailed record of my professional qualifications, including a list of patents and academic and professional publications, is set forth in my curriculum vitae attached to this declaration as Appendix 1.

8. I received a Ph.D. degree in Computer Science from the University of California at Berkeley in 2006. My Ph.D. Dissertation was entitled: "Lambda Data Grid: Communications Architecture in Support of Grid Computing."

9. I was granted a Master's of Science ("M.Sc.") degree in Electrical Engineering from Tel Aviv University, Israel in 1996.

10. I received a Bachelor of Science, ("B.Sc.") degree in Mathematics and Computer Science from Tel Aviv University, Israel in 1987.

11. I have over 25 years of experience in the networking, telecommunications, Internet, and software fields.

12. I currently am employed by the University of California at Berkeley and was appointed as a lecturer and an Industry Fellow in the Center of Entrepreneurship and Technology ("CET") as part of UC Berkeley College of Engineering.

13. I have been with the University of California at Berkeley since 2000 where I served as Berkeley Industry Fellow, Lecturer, Visiting Scientist, Ph.D. Candidate, and Nortel's Scientist Liaison. Some positions and projects were done concurrently, others sequentially.

14. I was appointed as a Principal Investigator for US Department of Defense (DARPA) Projects. For these projects, I conceived concepts, wrote proposals, and completed three research projects. In addition, I led a research

project for an undisclosed US Federal Agency. I led these projects for about 5 years while holding positions at Nortel Networks.

15. I have over 25 years of experience as a scientist, educator and technologist. I possess strong engineering background and ability to turn forward-looking academic research and novel concepts into products. I have been working mainly in research and advance technologies in the high-tech industry. My previous employers include Nortel Networks, Aptel Communications, Scitex and Shalev Robotics.

16. I am a Principal Scientist at my company Innovation IP, where I develop network communication technologies, provide research and consulting in advanced technologies, mainly in computer networking and Internet technologies. In this role, I bridge science, engineering and innovation to identify patentability. I analyze patents, build patent portfolios, and consult on the engineering and scientific aspects of patents.

17. I worked for Bay Networks and Nortel Networks for eleven years (Bay Networks was acquired by Nortel Networks). I held scientific and research roles at Nortel Labs, Bay Architecture Labs, and CTO Office in the fields of computer networking and Internet technologies. Positions included: Principal Scientist, Principal Architect, Principal Engineer, Senior Software Engineer.

18. I worked for Aptel communications for two years as a software engineer and a team leader. As part of my work, I developed Personal Communications Network (“PCN”) technologies.

19. I worked for Scitex Corporation for about four years as a software engineer and a team leader. Scitex was acquired by Hewlett Packard (“HP”). At Scitex, I worked on the networking and communications aspects of graphical applications for the pre-press industry.

20. I worked for Shalev Robotics for about three years, developing algorithms for robotics.

21. I am an advanced user of computer technologies for over 25 years, and during these years, I have been using leading-edge electronics, computers and Internet technologies.

22. I am named as a co-inventor on over 80 patents issued. I co-authored over 25 scientific publications, journal articles, and peer-reviewed papers. Furthermore, I'm a Senior Member of the Institute of Electrical and Electronics Engineers ("IEEE").

23. I have extensive experience in routing and switching architectures and protocols, including Multi-Protocol Label Switching Networks, Layer 2 and Layer 3 Virtual Private Networks, and Pseudowire technologies. I worked for Nortel Networks for over 11 years in research and development of these technologies. I wrote software for Bay Networks and Nortel Networks switches and routers. I developed network technologies for the Accelar 8600 switches and routers family, the OPTera 3500 SONET switches, the OPTera 5000 DWDM family, and for the Alteon L4-7 switching product family. I installed, configured and ran switches, routers and other network devices from Cisco Systems, Juniper Networks, Extreme Networks and other communication vendors.

24. I have a great deal of familiarity with the Internet Engineering Task Force ("IETF") and have closely followed the development of various networking standards protocols over the past 20 years. I had indirect exposure to the many IETF documents, presentations and strategies while working at Bay Architecture Lab and Nortel CTO Office. My direct peers were actively engaged in standard development work including IETF, IEEE, ITU and MEF. They heavily contributed to the standardization process, and I have personally reviewed many early and advanced drafts from these standards organizations.

25. I have extensive, personal, hands-on experience with technologies that deliver resiliency, priority, and preemption, as referred to by the '652 patent.

26. For example, the OPTera Metro 3500 Multiservice Platform is a commercial device that delivers the key features of the technology described in the '652 patent. The OPTera 3500 is a SONET switch providing advanced Ethernet services over metropolitan areas; it supports Resilient Packet Ring (RPR) and provides optical Ethernet private line (OE-PL) services using 10/100/1000 Ethernet. An OPTera 3500 specification from April 2004 (the time in which I worked on the product) is attached as Appendix 24.

27. I have extensive experience with the OPTera 3500, having written software for, installed, configured, used, written about and demonstrated the OPTera 3500. For example, I used the OPTera 3500 in a demonstration at a DARPA conference (held on May 29, 2002 in San Francisco, CA) and published a related paper (*DANCE 2002*, ISBN 0-7695-1564-9, IEEE Computer Society, p 344-354). Further, I used the OPTera 3500 as part of my presentation at two Supercomputing conferences: the first was in Phoenix, Arizona, from November 15-21, 2003; the second was in Pittsburg, Pennsylvania, from November 6-12, 2004.

28. I also presented several related presentations at UC Berkeley and other industry conferences and have published several related papers, as well.

#### **IV. BASIS FOR OPINION**

29. My opinions and views set forth in this declaration are based on my education, training, and experience in the relevant field, as well as the materials I reviewed in this case, and the scientific knowledge regarding the same subject matter that existed prior to the effective filing date of the '652 patent.

##### **A. Summary of Legal Principles**

30. In preparing my declaration and formulating my opinions, I have been provided the following summaries of some of the relevant legal principles. I am not a lawyer and do not intend to testify about legal issues, although I do have some familiarity with legal principles.

31. I understand that there are a number of legal factors or requirements that may be considered in determining whether the claims of a patent are valid or not. I also understand that, although the claims of an issued patent are presumed valid, those claims can be shown to be invalid by clear and convincing evidence that they fail to comply with one or more requirements of patentability. Notably, in situations where (as here) the Patent Office did not have all relevant information at its disposal during prosecution of the patents at issue, the considered judgment of the Patent Office in issuing those patents may lose significant force, i.e., the clear and convincing standard may be easier to sustain.

##### *1. Anticipation*

32. I understand that a person is not entitled to a patent if the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for patent. *See* 35 U.S.C. § 102(a).

33. I understand that a person is not entitled to a patent if the invention was patented or described in a printed publication in this or a foreign country or was in public use or on sale in this country more than one year prior to the date of



the application for patent in the United States. *See* 35 U.S.C. § 102(b). It is my further understanding that a sale or offer for sale may invalidate a patent under this section if what is sold or offered for sale is “ready for patenting,” i.e., it has either been reduced to practice or sufficient preparations have been made by the inventor to enable one of skill in the art to practice the invention. However, the parties to the transaction need not recognize that the product possesses the claimed characteristics.

34. I understand that a person is not entitled to a patent if the invention was described in a published application for a patent filed by another in the United States before the invention by the patent applicant, or a patent granted on an application for patent by another filed in the United States before the invention by the patent applicant. *See* 35 U.S.C. § 102(e).

35. I understand that a person is not entitled to a patent if he or she did not invent the subject matter sought to be patented (sometimes known as “derivation”). *See* 35 U.S.C. § 102(f).

36. I understand that a person is not entitled to a patent if before such person’s invention thereof, the invention was made in this country by another inventor who had not abandoned, suppressed, or concealed it. *See* 35 U.S.C. § 102(g)(2). In determining priority of invention, one must consider the respective dates of conception and reduction to practice of the invention, and also the reasonable diligence of one who was first to conceive and last to reduce to practice, from a time prior to conception by the other.

37. Consistent with these principles, I understand that a patent claim is invalid for a lack of “novelty” (also called “anticipation”) if what is claimed is not new. Anticipation occurs if, within the “four corners” of a single prior art reference, each and every limitation of the patent claim is disclosed, either explicitly or inherently. I have been informed that a claim limitation may be

inherently disclosed where it would have been necessarily present in the prior art device or method.

## 2. *Obviousness*

38. I understand that a patent claim may be found invalid as obvious if, at the time when the invention was made, the subject matter of the claim, considered as a whole, would have been obvious to a person having ordinary skill in the field of the technology (the “art”) to which the claimed subject matter belongs.

39. I understand that the following factors should be considered in analyzing obviousness: (1) the scope and content of the prior art; (2) the differences between the prior art and the claims; and (3) the level of ordinary skill in the pertinent art. I also understand that certain other factors known as “secondary considerations” such as commercial success, unexpected results, long felt but unsolved need, industry acclaim, simultaneous invention, copying by others, skepticism by experts in the field, and failure of others may be utilized as indicia of nonobviousness. I understand, however, that secondary considerations should be connected, or have a “nexus,” with the invention claimed in the patent at issue.

40. I understand that a person of ordinary skill in the art is assumed to have knowledge of all prior art. I understand that one skilled in the art can combine various prior art references based on the teachings of those prior art references, the general knowledge present in the art, or common sense. I understand that a motivation to combine references may be implicit in the prior art, and there is no requirement that there be an actual or explicit teaching to combine two references. Thus, one may take into account the inferences and creative steps that a person of ordinary skill in the art would employ to combine the known elements in the prior art in the manner claimed by the patent at issue. I understand that one should avoid “hindsight bias” and ex post reasoning in performing an

obviousness analysis. But this does not mean that a person of ordinary skill in the art, for purposes of the obviousness inquiry, does not have recourse to common sense.

41. I understand that when determining whether a patent claim is obvious in light of the prior art, neither the particular motivation for the patent nor the stated purpose of the patentee is controlling. The primary inquiry has to do with the objective reach of the claims, and that if those claims extend to something that is obvious, then the entire patent claim is invalid.

42. I understand one way that a patent can be found obvious is if there existed at the time of the invention a known problem for which there was an obvious solution encompassed by the patent's claims. I understand that a motivation to combine various prior art references to solve a particular problem may come from a variety of sources, including market demand or scientific literature. I understand that a need or problem known in the field at the time of the invention can also provide a reason to combine prior art references and render a patent claim invalid for obviousness.

43. I understand that familiar items may have obvious uses beyond their primary purpose, and that a person of ordinary skill in the art will be able to fit the teachings of multiple prior art references together "like the pieces of a puzzle." I understand that a person of ordinary skill is also a person of at least ordinary creativity.

44. I understand when there is a design need or market pressure to solve a problem and there are a finite number of identified, predictable solutions, a person of ordinary skill has good reason to pursue the known options within his or her technical grasp. If these finite number of predictable solutions lead to the anticipated success, I understand that the invention is likely the product of ordinary skill and common sense, and not of any sort of innovation. I understand that the

fact that a combination was obvious to try might also show that it was obvious, and hence invalid, under the patent laws.

45. I understand that if a patent claims a combination of familiar elements according to known methods, the combination is likely to be obvious when it does no more than yield predictable results. Thus, if a person of ordinary skill in the art can implement a predictable variation, an invention is likely obvious. I understand that combining embodiments disclosed near each other in a prior art reference would not ordinarily require a leap of inventiveness.

**B. A Person of Ordinary Skill in the Art**

46. It is my opinion that a person of ordinary skill in the art with respect to the '652 patent as of 2005-2006 would have a bachelor's degree in computer science, electrical engineering or the equivalent thereof and at least 7 years of professional experience within the field of network communications and internet protocols; or an advanced degree in in computer science, electrical engineering or the equivalent thereof and at least 4 years of professional experience within the field of network communications and internet protocols.

47. A person of ordinary skill in the art would also be familiar with the development of industry standards and protocols, including familiarity with the Internet Engineering Task Force ("IETF"). He or she would be aware of standard protocols for the establishment and maintenance of Label-Switched Paths ("LSP"), Pseudowires ("PW") and other virtual paths, including the Label Distribution Protocol ("LDP"), Resource Reservation Protocol ("RSVP"), and other Traffic Engineering ("TE") protocols used to set up, control and manage virtual paths in MPLS, IP, and hybrid Layer 2/Layer 3 networks.

48. A network communications device does not function in a vacuum; it must interact and interoperate with other switches, routers, gateways, and network devices. Therefore, those with familiarity of network communications standards

and protocols have a profound knowledge of the software, hardware, design, architecture, and interoperability of products not only from their own companies but also from those of other major, mid-range, and minor contributors to the field.

49. Indeed, network communications products must work correctly in a variety of environments; with a variety of other products from a variety of other companies; across multiple service providers; with various autonomous systems; and regardless of vendor. Due to the necessity of complex and broad-ranging interoperability, a person with knowledge of the relevant industry standards must have a broad and deep understanding of the field.

50. My opinions regarding the level of ordinary skill in the art are based on, among other things, my over 25 years of experience in the field of network communications, computer science and engineering, my understanding of the basic qualifications that would be relevant to an engineer or scientist tasked with investigating methods and systems in the relevant area, and my familiarity with the backgrounds of colleagues and co-workers, both past and present.

## V. THE '652 PATENT

### A. Overview

51. The claims of the '652 patent are directed to a technique for configuring and using a “standby Pseudowire” that involves assigning a “priority” to a standby Pseudowire during configuration and then using that “priority” to determine whether to “preempt existing traffic on the standby Pseudowire.”

52. The patent defines a “Pseudowire” as an emulation of a native service (such as Asynchronous Transfer Mode (“ATM”), Ethernet, Time Division Multiplexing (“TDM”), Synchronous Optical Network (“SONET”), or Synchronous Digital Hierarchy (“SDH”)) over a packet-switched network (such as Multiprotocol Label Switching (“MPLS”) or Internet Protocol (“IP”)). *Id.* at 1:14-25.

53. Claim 1 is exemplary:

1. A method of providing protection to network traffic, comprising:
  - sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;
  - receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;
  - accepting the Pseudowire protection configuration parameter by the destination node;
  - using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and
  - determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

54. Independent claims 9 and 14 are similar to claim 1, except that they claim a system and a computer program product, respectively. Thus, the challenged independent claims relate generally to signaling a standby Pseudowire by sending a configuration parameter that contains a protection property that includes a priority for the standby Pseudowire between a source node and a destination node, receiving an acknowledgement from the destination node indicating that the requested parameters have been accepted, using the standby Pseudowire, and then determining whether to preempt existing traffic on the standby Pseudowire based, at least in part, on the priority for the standby Pseudowire.

55. The challenged dependent claims are related to (1) providing protection to at least one primary Pseudowire (claims 2 and 10), (2) switching network traffic from a primary Pseudowire to the standby Pseudowire when there is a network failure (claim 3), (3) dynamically selecting the standby Pseudowire from a plurality of connections (claim 4), (4) including a domain type, protection type of protection scheme in the configuration parameter (in addition to the priority) (claims 5, 11, and 15); and (5) having a protection scheme that indicates a 1+1, 1:1, 1:N, or M:N protection scheme (claims 8, 13 and 17).

56. Collectively, I refer to claims 1-5, 8-11, 13-15, and 17 as the “Challenged Claims.”

#### **B. Priority of the '652 Patent**

57. I have been informed that, for a claim to benefit from the earlier filing date of a provisional application, each element of the claim must be disclosed in the provisional application so that someone of ordinary skill in the art would be able to make and use the claimed invention.

58. The '652 patent was filed on February 14, 2006 and claims priority to Provisional Application No. 60/653,065, which was filed on February 14, 2005.

The specification of the '652 patent varies dramatically from that of the provisional application.

59. I understand that BSL has not produced any evidence of conception or reduction to practice that pre-dates the provisional filing date of the '652 patent. Because the prior art that I cite in this declaration pre-dates the provisional applications, I have not undertaken a detailed analysis regarding whether each of the Challenged Claims is entitled to the provisional filing date at this time. To the extent that BSL later attempts to swear behind any of the references cited in this declaration or in Juniper's Petition, I reserve the right to supplement my declaration to address those arguments.

### **C. Overview and Background of the Technology**

60. Ping Pan did not invent Pseudowires; nor did he invent the concept of using configuration parameters (e.g., modes, priorities, protection schemes, etc.) to provide protection to traffic on a data network. Pan also did not invent the concept of using priorities to make decisions about traffic preemption.

61. Rather, as I explain below, each of the main elements of the '652 patent existed and was well-known in the prior art long before February 2005. Moreover, the prior art shows that market trends and pressures in this area would naturally suggest to one of skill in the art the very protection techniques described by the '652 patent. As a result of these trends and forces, it would have been obvious to combine the various prior art references, as described below.

62. The technology discussed in the '652 patent is applied in a very narrow field (protective internet protocols in a Pseudowire environment (*i.e.*, a hybrid Layer 2/Layer 3 network), in the field of network communications). As shown below, within this narrow field, the concept of protecting Pseudowires was a mainstream, expected method of configuring routers and switches; prior to 2005, it was both well-published—for example, in various standards setting organization



documents and industry books—and commonly configured in products in the market.

1. *Internet Engineering Task Force (IETF)*

63. The Internet Engineering Task Force (IETF) is an open organization that develops and promotes networking and Internet standards. By 2000, all of the major industry players—including network product providers such as Cisco, Juniper, Nortel, and Tellabs, and service providers such as AT&T and Level 3 Communications—as well as smaller Internet technology companies, were participating in the IETF. The IETF served as a forum for industry collaboration to standardize Internet technologies.

64. To facilitate this process, members of the IETF submitted “Internet Drafts” to propose interoperability solutions and standards for a variety of Internet technologies. The drafts were reviewed, vetted, presented at IETF conferences, and refined by the members of the group. After vetting, some Internet Drafts are published as “Requests for Comments” (“RFCs”). The IETF community also publishes other types of documents such as Technical Specifications, Applicability Statements, Proposed Standards, Draft Standards, and Internet Standards. These documents are published on IETF’s discussion boards and circulated to large distribution mailing lists.

65. The proposed standards developed by these groups were widely circulated and well-known to those skilled in the art. For example, they were published on the IETF website and circulated to members in the form of Internet Drafts or RFCs.

66. During the time period from 2000-2005, there were several collaborative groups organized by the IETF that were actively involved in setting standards and protocols for the particular technologies that are relevant to the '652 patent. For example, the IETF had chartered working groups to address the

standardization of protocols related to Multi-Protocol Label Switching (“MPLS”), Traffic Engineering, Layer 2 and Layer 3 Virtual Private Networks (“VPN”), and Pseudowires (sometimes referred to as “PWE3”).

67. Due to the nature of this area—in which a network communications device must interact and interoperate with other switches, routers, gateways, and network devices—those who set network communications standards had a profound knowledge of the software, hardware, design, architecture, and interoperability of the products.

68. As an active member of the industry who was working on the development of network switches and routers, I was well-aware of the emerging standards in this area. Indeed, compatibility and the ability to comply with industry standards is an important feature of any switch or router.

69. The activities of the IETF are highly relevant to the ’652 patent. In fact, based on my review of the IETF archive of Internet Drafts, it appears that the inventor of the ’652 submitted various Internet Drafts to the IETF community that are closely related to the concepts discussed in the ’652 patent, several of which appear to be collaborations with other IETF members. *See, e.g.*, App. 28 (Ping Pan IETF Drafts). It also appears that Pan presented and/or co-presented at several IETF conferences on topics that are similar to those discussed in the ’652 patent. *See, e.g.*, App. 29 (Ping Pan IETF Presentations).

## 2. *Packet-Switched Networks and MPLS*

70. By way of background, “Packet-Switched Networks” are a type of communications network that route data through a network in small units of data called “packets.” Packets typically have a “header” that provides relevant information about the packaged data so that the nodes on the network can properly route it. Most traffic over the Internet uses packets.

71. In a traditional packet-switched network, each router on the network examines the packet's header to make decisions about where to route the packet. As a result, every time a packet arrives at a router, the router needs to "unpack" the packet and analyze it to determine where it should be sent next.

72. Multi-Protocol Label Switching ("MPLS") is a technique that can be used in lieu of traditional packet switching. In a network that uses MPLS, a packet that enters the network is assigned to a specific forwarding equivalence class ("FEC"). The FEC can be based on various attributes of the packet, such as the particular port it comes from or the type of application. The FEC is indicated by adding a "label" to the packet that consists of a short bit sequence. Each router in the MPLS network has a table that tells the router what to do with packets that have been designated as belonging to a specific FEC. Thus, subsequent routers in an MPLS network need not "unpack" the packet and analyze it to determine where it should be sent. Instead, the subsequent routers use the label to look up the appropriate entry in the table and provide the packet with a new FEC according to the rules in the table. This allows the routers in an MPLS network to route in higher bandwidth and in a more efficient manner.

73. The paths or "tunnels" in a standard MPLS network are typically called Label-Switched Paths or "LSPs."

### 3. *Pseudowires*

74. The '652 patent defines a Pseudowire as an emulation of a native service (such as Asynchronous Transfer Mode ("ATM"), Ethernet, Time Division Multiplexing ("TDM"), Synchronous Optical Network ("SONET"), or Synchronous Digital Hierarchy ("SDH")) over a packet-switched network (such as Multiprotocol Label Switching ("MPLS") or Internet Protocol ("IP")). *Id.* at 1:14-25. I understand that BSL has taken a broader view of "Pseudowire," defining it as

any “emulation of a native service over a network.” App. 26 (Joint Claim Construction Chart) at Exhibit A, page 3.

75. During the time frame starting in 2000, there were several emerging techniques coming out of the IETF community for emulating a native service (such as ATM, Ethernet, TDM, SONET, or SDH) over packet-switched MPLS or IP networks.

76. On a high-level, these techniques involved encapsulating the native data into a packet and using additional label fields to represent the relevant native characteristics of the data, so that the data could be sent like a packet over the MPLS or IP network while still maintaining all the necessary characteristics once reaching its destination.

77. As one example, in mid-2000, a group of engineers from Cisco Systems and Level 3 Communications (including Luca Martini, Eric Rosen, and Nasser El-Aawar, among others), began developing a specification called “Encapsulation Methods for Transport of Layer 2 Frames over MPLS Networks” that ultimately became known as the “draft-martini” protocol. The “draft-martini” protocol described a detailed method of encapsulation that allowed for the emulation of native services over an MPLS network. In essence, “draft-martini” described a way to set up “emulated virtual circuits” to carry the Protocol Data Units (“PDUs”) used in the various Layer 2 protocols (*e.g.*, Ethernet, ATM, etc.) across the MPLS network. “Draft-martini” provided various methods for the preserving the PDUs, such as using a “control word” and/or adding a “VC label” to the MPLS label stack.

78. Through additional work and collaboration in the IETF community, “draft-martini” protocol was refined into a subsequent specification by an overlapping group of authors called “Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP),” the first draft of which was published on the

IETF website in August 2002. This draft ultimately became RFC 4447, and provides additional details regarding the types of labels that are necessary to transport Layer 2 packets over an MPLS network. RFC 4447 also discusses the use of the traditional Label Distribution Protocol signaling (discussed further below) to establish Pseudowires.

79. As another example, a separate specification called “Virtual Private LAN Services over MPLS (VPLS)” was introduced to the IETF community in 2003 by a large group of engineers from numerous companies, led by Marc Lasserre and Vach Kompella (“Lasserre Draft”). The Lasserre Draft describes extensions to the Pseudowire protocols developed by Martini and his group that allow for the transport of Ethernet and VLAN traffic across multiple sites that belong to the same Layer 2 broadcast domain or VPLS. App. 19 (Lasserre Draft) Version 1 at Abstract, § 4.

#### 4. *MPLS Network Signaling Protocols*

80. LSPs, Pseudowires, and other types of virtual paths in an MPLS network are typically set up between routers or other nodes in an MPLS network using a standardized signaling protocol. By early 2005, prior to the ’652 patent, there were several conventional signaling protocols that were widely accepted and used for this purpose.

81. One example is the Label Distribution Protocol (“LDP”). LDP was a precursor technology to the Pseudowire concept discussed above. It is described in a number of RFCs and IETF drafts, including RFC 3036, which was published in January 2001. As described in RFC 3036, LDP works by having a Label-Switched Router (“LSR”) send a “Label Request” to another LSR to initiate the request to set up an LDP tunnel. App. 10 (RFC 3036) at § 3.5.8. The Label Request contains the relevant FEC, along with various other configuration parameters. *Id.* at 3.5.8. The “value” part of a TLV may itself contain one or more additional TLVs. *Id.* at

1.3. Once the receiving LSR has processed the request, it either responds with a Label Mapping for the requested label or with a Notification message indicating why it cannot satisfy the request. *Id.* at 3.5.8.1. All LDP messages have a common structure that uses a Type-Length-Value (TLV) encoding scheme. Once this process is completed, the LSP, Pseudowire or other virtual tunnel is set up and can be used to transmit traffic according to the parameters set forth in the Label Request.

82. Another example of a signaling protocol that is used to set up LSPs, Pseudowires, and other LSPs in an MPLS network is the Resource Reservation Protocol or “RSVP.” RSVP was also extended to “RSVP-TE” (i.e., RSVP-Traffic Engineering), which is described in a number of IETF Drafts and RFCs and which provides additional mechanisms for traffic engineering (i.e., more control and management of network tunnels during set-up and operation). For example, RFC 3209, was published by the IETF in December of 2001. In this protocol, a “sender node” sends a “Path Message” with a “LABEL\_REQUEST object” to a “destination node.” App. 9 (RFC 3209) at § 2.2 (“Operation of LSP Tunnels”). The Path Message contains information about the requested parameters for the LSP or other virtual path that it would like to set up. *Id.* The destination node of a label-switched path responds to a LABEL\_REQUEST by including a LABEL object in its response, which is called an RSVP Resv message. *Id.* RSVP-TE protocol identifies a number of specific “objects” that can be used during signaling to give the service provider more control over the virtual paths, such as setup/holding priorities and protection styles (which will be discussed in more detail below). *Id.* Once this exchange is completed, the path is ready to be used according to the specified configuration parameters.

83. Thus, the standardized protocols for setting up LSPs, Pseudowires, and Tunnels in an MPLS network involved (1) a source node sending a message to

a destination node that includes a request to set up the path, along with the particular configuration parameters that are required for the path, (2) the destination node processing the request and responding to the source node by indicating whether the parameters have been accepted, (3) using the path according to the specified parameters.

#### 5. *Network Protection, Survivability and Resiliency*

84. Dropped packets and link failure in communications networks were and are well-known concepts in the field; all networks have a finite capacity; the quantity of traffic often exceeds the capacity of a network. To meet customer needs and expectations, an important aspect of any data communications network is the protection of traffic on the network.

85. Therefore, network service providers deal with these issues using the **concept of priority**. Priority is a technique used to determine which network packets will go through and which will be **preempted (dropped)**. Networks employ the concepts of priority and preemption by the use of service layer agreements (SLA), quality of service (QoS), policy, queuing, and/or other similar policies to prioritize and preempt traffic. For example, to the extent a customer needs a higher degree of protection; service providers can use these features to prioritize that customer's traffic in the event of a failure network overload.

86. More specifically, a service level agreement (SLA) is an agreement that a network service provider has with its customers about type of service it will give them. Some customers chose to pay more to get better service. Some chose to pay less and receive lower quality service. The customers who pay more for SLA have fewer dropped packets (for example, in cases of congestion) because they have higher priority. In case of line failure, the customer with a better SLA will get higher priority on alternative routes.

87. As an example, if we assume there are only two types of traffic: one higher priority type of traffic (for which a customer has paid the network service provider more money for a better quality of service (QoS)) and a second lower priority type of traffic (for which a customer has paid less for a lower QoS). All other things being equal, the traffic with the better QoS will get better network service and will be considered higher priority.

88. Another similar concept is “policy.” Policy is the condition(s) in which variable priority (or service level) is given to different customers, applications, and types of traffic. The architecture of network traffic is based on queuing. Packets metaphorically “stand in line”; those with a higher priority on the switch will, metaphorically, move to the head of the line and leave the queue faster.

89. Network communication systems often intentionally drop some packets and allow others to slow down. For example, Server TCP mechanisms intentionally drop off packets to signal the application to slow down. The server would only drop packets from some types of traffic (e.g., lower priority) but not others to maximize resources.

90. SLA, priority, queuing, QoS, class of service, classification, policy, dropping, and packet dropping are all basic concepts of network communication. They are fundamentals that any student learns in his or her first introductory network communications class. Certainly, any network engineer knows how to design a network to handle failure.

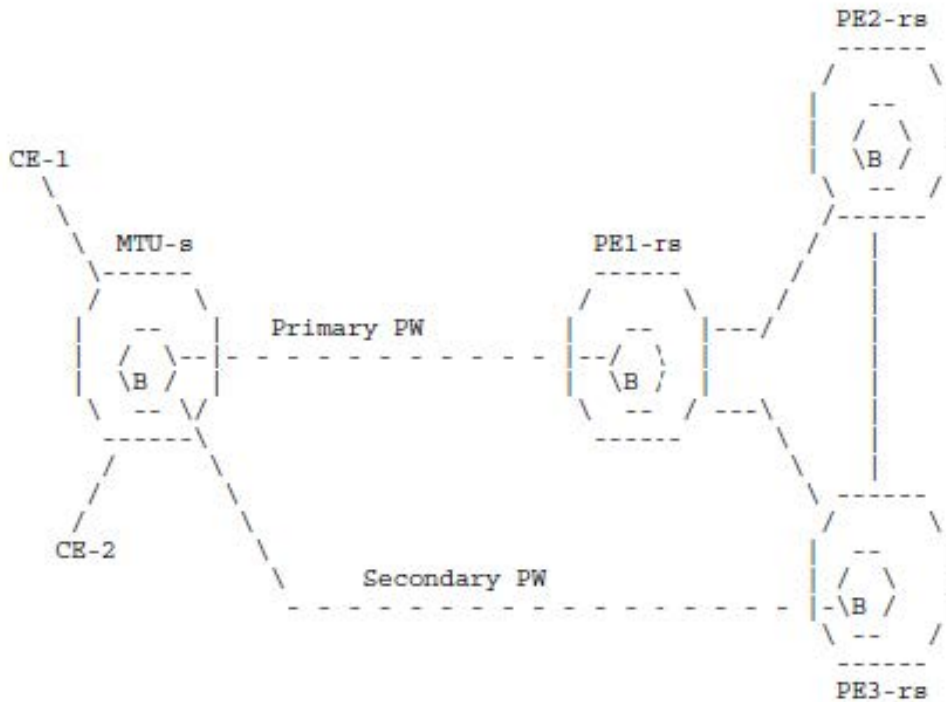
91. In the specific context of MPLS and Pseudowires, prior to the '652 patent, there were a plethora of well-known methods for protecting traffic in MPLS and Pseudowire networks. Some examples are provided below.

92. For example, the Lasserre Draft mentioned above describes basic redundancy and traffic protection features that can be employed in a Pseudowire



environment. For example, Lasserre teaches that a “primary PW” can be protected by a “secondary PW” using “Dual-homed MTU devices.” Lasserre describes and depicts this technique as follows:

An MTU-s device will setup two [PWE3-ETHERNET] pseudowires (one each to PE-rs1 and PE-rs2) for each VPLS instance. One of the two pseudowires is designated as primary and is the one that is actively used under normal conditions, while the second pseudowire is designated as secondary and is held in a standby state. The MTU device negotiates the pseudowire labels for both the primary and secondary pseudowires, but does not use the secondary pseudowire unless the primary pseudowire fails. Since only one link is active at a given time, a loop does not exist and hence 802.1D spanning tree is not required.



*Id.* at § 10.2.1

93. In addition to the protection scheme described in Lasserre, there were numerous disclosures of more sophisticated protection features for LSPs and Pseudowires in an MPLS network.

94. For example, the concept of assigning a “Setup Priority” and/or “Holding Priority” had been disclosed in numerous RFCs and patents. By way of example, RFC 3209 discloses that, a “SESSION\_Attribute object can be added to Path messages to aid in session identification and diagnostics. Additional control information, such as setup and hold priorities, resource affinities, . . . and local-protection, are also included in this object. App. 9 (RFC 3209) at § 2.2. RFC 3209 further explains that the “Setup Priority is used in deciding whether this session can preempt another session” and the “Holding Priority is used in deciding whether this session can be preempted by another session.” *Id.* at § 4.7.1. RFC 3209 further shows how these priorities are included in the SESSION\_Attribute object:

4.7. Session Attribute Object

The Session Attribute Class is 207. Two C\_Types are defined, LSP\_TUNNEL, C-Type = 7 and LSP\_TUNNEL\_RA, C-Type = 1. The LSP\_TUNNEL\_RA C-Type includes all the same fields as the LSP\_TUNNEL C-Type. Additionally it carries resource affinity information. The formats are as follows:

4.7.1. Format without resource affinities

```
SESSION_ATTRIBUTE class = 207, LSP_TUNNEL C-Type = 7

  0                               1                               2                               3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Setup Prio | Holding Prio |      Flags      | Name Length |
+-----+-----+-----+-----+-----+-----+-----+-----+
//           Session Name           (NULL padded display string)           //
|-----|-----|-----|-----|-----|-----|-----|-----|
```

*Id.* at §§ 4.7 and 4.7.1

95. The topic of Setup Priorities and Holding Priorities is also discussed in RFC 4090 (Appendix 14) and U.S. Patent Application Publication No. 2004/0156313 (August 12, 2004) to Ralph Theodore Hofmeister and Ping Pan (Appendix 4).

96. As another example, RFC 3386, entitled “Network Hierarchy and Multilayer Survivability” was published by the IETF in November 2002. It describes various configuration parameters that can be used in connection with traditional signaling protocols to provide traffic protection in a wide range of networks, including SONET, MPLS, GMPLS and Pseudowire environments. App. 5 (RFC 3386) at § 1.

97. RFC 3386 teaches that a “working entity” can be used to “carry traffic in normal operation mode” and, depending on the context, it can be a channel or link, a path, or a bundle of paths. *Id.* at § 2.2.2. RFC 3386 further teaches that a “protection entity” (otherwise called “backup entity” or “recovery entity”) can be “used to carry protected traffic in recovery operation mode, i.e., when the working entity is in error or has failed.” *Id.* RFC 3386 further teaches that “extra traffic” or “preemptable traffic” can be “carried over the protection entity while the working entity is active. Extra traffic is not protected, i.e., when the protection entity is required to protect the traffic that is being carried over the working entity, the extra traffic is preempted.” *Id.*

98. RFC 3386 goes on to discuss a number of different recovery modes, such as revertive and non-revertive. In revertive mode, the protected traffic is switched back from the protection entity to the working entity once the failed working entity has been prepared, whereas in non-revertive mode there is no “preferred path” and the backup path becomes the new “working” entity. *Id.*

99. RFC 3386 also describes various survivability techniques at length. For example, RFC 3386 teaches that “[p]rotection techniques can be implemented

by several architectures: 1+1, 1:n, 1:n, and m:n.” *Id.* at § 2.2.3. RFC 3386 explains these techniques as follows:

In the 1+1 protection architecture, a protection entity is dedicated to each working entity. The dual-feed mechanism is used whereby the working entity is permanently bridged onto the protection entity at the source of the protected domain. In normal operation mode, identical traffic is transmitted simultaneously on both the working and protection entities. At the other end (sink) of the protected domain, both feeds are monitored for alarms and maintenance signals. A selection between the working and protection entity is made based on some predetermined criteria, such as the transmission performance requirements or defect indication.

In the 1:1 protection architecture, a protection entity is also dedicated to each working entity. The protected traffic is normally transmitted by the working entity. When the working entity fails, the protected traffic is switched to the protection entity. The two ends of the protected domain must signal detection of the fault and initiate the switchover.

In the 1:n protection architecture, a dedicated protection entity is shared by n working entities. In this case, not all of the affected traffic may be protected.

The m:n architecture is a generalization of the 1:n architecture. Typically  $m \leq n$ , where m dedicated protection entities are shared by n working entities.

*Id.*

100. RFC 3386 further describes the use of “restoration priority” to “giv[e] preference to protect higher-priority traffic ahead of lower-priority traffic.” *Id.* at § 2.2.4. RFC 3386 also describes the use of “preemption priority” which is “a method of determining which traffic can be disconnected in the event that not all traffic with a higher restoration priority is restored after the occurrence of a failure.” *Id.*

101. In light of these concepts, RFC 3386 discusses the pros and cons of a “1+1” architecture vs. “1:1” vs. “1:n” vs. “m:n”:

A 1+1 protection architecture is rather expensive since resource duplication is required for the working and protection entities. It is generally used for specific services that need a very high availability.

A 1:1 architecture is inherently slower in recovering from failure than a 1+1 architecture since communication between both ends of the protection domain is required to perform the switch-over operation. An advantage is that the protection entity can optionally be used to carry low-priority extra traffic in normal operation, if traffic preemption is allowed. Packet networks can pre-establish a protection path for later use with pre-planned but not pre-reserved capacity. That is, if no packets are sent onto a protection path,

then no bandwidth is consumed. This is not the case in transmission networks like optical or TDM where path establishment and resource reservation cannot be decoupled.

In the 1:n protection architecture, traffic is normally sent on the working entities. When multiple working entities have failed simultaneously, only one of them can be restored by the common protection entity. This contention could be resolved by assigning a different preemptive priority to each working entity. As in the 1:1 case, the protection entity can optionally be used to carry preemptable traffic in normal operation.

While the m:n architecture can improve system availability with small cost increases, it has rarely been implemented or standardized.

*Id.* at § 2.3

102. RFC also describes how a protection path can be set up using traditional signaling protocols and configuration parameters:

### 3.2 Required initial set of survivability mechanisms

#### 3.2.1 1:1 Path Protection with Pre-Established Capacity

In this protection mode, the head end of a working connection establishes a protection connection to the destination. There should be the ability to maintain relative restoration priorities between working and protection connections, as well as between different classes of protection connections.

In normal operation, traffic is only sent on the working connection, though the ability to signal that traffic will be sent on both connections (1+1 Path for signaling purposes) would be valuable in non-packet networks. Some distinction between working and protection connections is likely, either through explicit objects, or preferably through implicit methods such as general classes or priorities. Head ends need the ability to create connections that are as failure disjoint as possible from each other. This requires SRG information

that can be generally assigned to either nodes or links and propagated through the control or management plane. In this mechanism, capacity in the protection connection is pre-established, however it should be capable of carrying preemptable extra traffic in non-packet networks. When protection capacity is called into service during recovery, there should be the ability to promote the protection connection to working status (for non-revertive mode operation) with some form of make-before-break capability.

*Id.* at §§ 3.2, 3.2.1. Thus, RFC 3386 teaches all of the key concepts of the '652 patent: Pseudowires, signaling using configuration parameters, assigning a priority, and preempting traffic based on a priority.

103. The protection techniques discussed in RFC 3386 are also disclosed in a number of other RFCs and patents.

104. For example, U.S. Patent No. 7,804,767 B1 to Kenneth Owens, Srinivas Makam, Changeheng Huang, and Vishal Sharma of Tellabs Operations, Inc. ("Owens"), which was filed in October 25, 2000, discloses the techniques described in RFC 3386 in the specific context of an MPLS network. *See* App. 15 (Owens). Owens describes a number of different protection-related configuration parameters that can be deployed in an MPLS environment. For example, Owens

notes that protection can be based on dynamically created paths at the time of failure or pre-negotiated protection paths. *Id.* at 5:1-29. Owens also describes various protection modes (e.g., revertive or non-revertive) and protection switching options (e.g., 1+1, 1:1, 1:n, and n:m). *Id.* at 6:16 – 7:15. As with RFC 3386, Owens notes that, in 1+1 protection, the protection path “could be used to transmitted an exact copy of the working traffic, with a selection between the traffic on the working and protection paths being made at the [destination].” *Id.* at 6:55-58. And, Owens also notes that in 1:1 protection, “the working traffic normally travels only on the working path, and is switched to the protection path only when the working entity is unavailable. Once the protection switch is initiated, all the low priority traffic being carried on the protection path is discarded to free resources for the working traffic.” *Id.* at 7:1-6.

105. Thus, both RFC 3386 and Owens contain detailed disclosures regarding the use of priorities, as well as the technique of preempting traffic based on those priorities, in the event of a failure.

106. The concepts of priority and preemption were discussed in myriad other IETF, IEEE, and Metro Ethernet Forum documents. Indeed, in November 2000 the IEEE began working on the 802.17 standard (known as Resilient Packet Ring or RPR) specifically to standardize the technology among the different players in the field. The standard was published on June 2004, well before the priority date of the '652 patent. This IEEE standard involved assigning traffic a CoS (Class of Service) that was used to prioritize certain traffic.

107. The purpose of the MEF (Metro Ethernet Forum) was to drive this known technology, market it, and demonstrate its capabilities and interoperability across the industry.

6. *Cisco Press Books*

108. In addition to being fully disclosed in a number of IETF documents and patents, the key aspects of the '652 patent were also summarized in a number of books published by Cisco Press.

109. The Cisco Press books are educational books that often summarize the state of the art and the work in the industry over the previous 3-5 years regarding the particular topic. The Cisco Press books are mainstream references that summarize hundreds of articles, papers, industry standards, public presentations and other reference materials.

110. The Cisco Press books were normal preparation for Cisco certification exams and other technical courses and provide multiple specific examples of configuring routers and switches to accomplish the networking topics described therein.

111. The Cisco Press books were also used by network engineers who were working with Cisco products or who wanted to learn about the capabilities of Cisco products. The Cisco Press books also often had content that is generally applicable to networking products, including products from the other leading network product providers such as Juniper, Alcatel and Nortel.

112. Thus the concepts in the Cisco Press book series generally are not unique; rather, they are expected knowledge among network engineers configuring Cisco routers and switches, as well as other popular network products.

113. One of the Cisco Press books that fully discloses each element of the '652 patent is called "Metro Ethernet" by Sam Halabi ("Halabi"). App. 6 (Halabi). It was published by Cisco Press on October 1, 2003, long before the '652 patent was filed.

114. Halabi discusses the adoption of Metro Ethernet services, as well as how those services have led carriers to deliver Metro Data Services. The book



delves into the role of virtual private networks (VPN), virtual private local area networks (VLAN), virtual private LAN services (VPLS), traffic engineering, and MPLS and Generalized MPLS (GMPLS) in the Metro Ethernet.

115. More specifically, the book examines the concepts of Virtual Private LAN Service (VPLS), SONET/SDH, Resilient Packet Ring (RPR), Pseudowire concept, Pseudowire via Layer 2 Tunneling Protocol (L2TP), Ethernet transport, and Ethernet over MPLS (“Draft Martini”).

116. It also covers various issues pertaining to the configuration and protection of hybrid Layer 2/Layer 3 IP/MPLS networks, along with the emulation of Layer 2 Ethernet services over MPLS networks, and the emulation of Layer 2 VPN over an IP network. This emulation of native services over a packet-switched network is also referred to as a “Pseudowire” environment.

117. Halabi also contains specific chapters that cover the concepts of RSVP-TE (RSVP signaling for traffic engineering) and fast MPLS Fast-Reroute, which were well-known techniques that allowed for greater control of network set-up and operation (as described above).

118. Halabi also covers the topic of Generalized MPLS (GMPLS), which is a protocol that allows various additional network resources (e.g., SONET, SDH, DWDM, and optical fibers) to be sent over an MPLS-like backbone.

119. Halabi is rich with information regarding MPLS, Pseudowires, and various protection techniques. Indeed, the book is a summary of approximately 3-5 years of mainstream knowledge in the field, referencing RFC documents, industry publications, Metro Ethernet Forum, IETF, ITU, and ETSI references related to Metro Ethernet.

120. Because Halabi, like other Cisco Press books, was normal preparation for Cisco certification exams and provides multiple specific examples of configuring routers and switches to accomplish the claimed ’652 invention, it

follows that the material was widely used by installing and configuring commercial Cisco switches and routers as well as by other competitors' products in the industry

121. The various discussions in Halabi fully and thoroughly disclose each and every element of the Challenged Claims of the '652 patent in detail.

122. Thus, the concepts detailed in Halabi were not unique; indeed, it was expected knowledge among the network engineers configuring Cisco routers and switches.

123. Similar concepts are also discussed in various other Cisco Press books, such as "Layer 2 VPN Architectures" by Wei Luo, Carlos Pignataro, Dmitry Bokotey and Anthony Chan (Appendix 22) and "Internet Routing Architectures," Second Edition, by Sam Halabi with Danny McPherson (Appendix 23).

#### **D. Prosecution History of the '652 Patent**

124. United States Patent Application No.11/354,569, which ultimately became the '652 patent was submitted on February 14, 2006. It was rejected numerous times prior to a Panel Decision for its allowance.

125. The originally-filed independent claims of the '652 patent required only four elements: (1) sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a designation node, (2) receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted, (3) in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire, and (4) wherein the standby Pseudowire is configured based at least in part on the Pseudowire configuration parameter. App. 3 ('652 File History) at 322.

126. Various dependent claims, including original claims 8, 9, 16, and 21 added the requirements that (1) the protection configuration parameter include a

priority (claims 8, 16 and 21) and, (2) determining whether to preempt existing traffic on the standby Pseudowire, the determination being based at least in part on a priority associated with the standby Pseudowire (claim 9). *Id.* at 322-324.

127. On November 20, 2008, the Examiner rejected the original claims as being unpatentable over Huang, 2003/017950 in view of the admitted prior art and in view of Blanchet, U.S. 2004/0133692. *Id.* at 265-279. The Examiner found that Huang disclosed the elements of sending a configuration parameter for configuring a standby path between a source node and a destination node, and using the standby path. *Id.* at 268-269. The Examiner further found that while Huang did not expressly teach Pseudowires, it would have been obvious to implement Pseudowires as a type of network service in the Huang system over the admitted prior art because “Pseudowires can emulate the operation of a ‘transparent wire’ carrying the native service” and the “method of modifying the system of Huang was within the ordinary ability of one of ordinary skill in the art based on the teachings of [the admitted prior art].” *Id.* at 269-270. The Examiner also found that Huang did not teach a configuration acknowledgement, but that Blanchet disclosed this element and that it would have been obvious to modify the Huang system to send an ACK message so as to make the system more reliable. *Id.* at 270. The Examiner also rejected original dependent claims 8-9, 16 and 21 under § 103(a) as being unpatentable in further view of Saleh, U.S. Patent No. 7,200,104, which teaches a priority, as well as determining whether to preempt existing traffic on a standby path based at least in part on the priority. *Id.* at 277-278.

128. In February of 2009, the applicant amended the independent claims to add the requirement that the Pseudowire protection configuration parameter indicate a protection property associated with the standby Pseudowire, and argued that the Examiner’s obviousness arguments were unwarranted in light of the

secondary considerations of nonobviousness discussed in the Background section of the specification. *Id.* at 258, 261-263.

129. On June 22, 2009, the Examiner again rejected the claims in a Final Office Action, this time under § 103 based on Huang in view of Voit, U.S. App. No. 2006/0047851, Blanchet, and Sridhar, U.S. App. No. 2006/0018252. *Id.* at 220-234. In addition to the elements disclosed by Huang and Blanchet, the Examiner found that Voit teaches Pseudowire protection and Sridhar teaches a configuration parameter that indicates a protection property associated with a standby link. *Id.* at 224. The Examiner further found that it would have been obvious to one skilled in the art to combine the system of Huang with Voit, Blanchet, and Sridhar to increase efficiency and solve the problem of data traffic protection because all of the references are in the same field of endeavor—network transfer. *Id.* As to the dependent claims involving priority and preemption, the Examiner found that Saleh taught these elements.

130. On December 22, 2009, the patentee filed a Request for Continued Examination, along with a Preliminary Amendment, which made several non-substantive amendments to the claims. *Id.* at 208-213.

131. On January 22, 2010, the Examiner issued Non-Final Office Action, rejecting the claims for the same reasons stated in the June 22, 2009 Office Action. *Id.* at pgs. 168-181. The Examiner further found that the applicant's arguments concerning non-obviousness were not persuasive in light of the fact that Voit teaches Pseudowire and Pseudowire protection, thus demonstrating that Pseudowire protection was not a new concept. *Id.* at pg. 170.

132. On April 21, 2010, the applicant again amended the claims, this time adding the limitations of dependent claims 8-9, 16 and 21 – that the protection property include a “priority” and the priority be used to determine whether to “preempt existing traffic on the standby Pseudowire” – to the independent claims.

*Id.* at pg. 145. The applicant further argued that Saleh (which the Examiner had cited in a previous Office Action) does not disclose these limitations because the “priority” disclosed by Saleh is used for selecting which nodes can be part of a virtual path (QoS) and for prioritizing a virtual path (which consists of both a primary and secondary path), and there thus is no separate priority for the primary vs. secondary paths. *Id.* at pgs. 151-152. The applicant also argued that the 1+1 protection scheme in Saleh did not disclose preempting “*existing* traffic” on the standby path because the secondary path of Saleh “is dedicated to the given VP for restoration purposes and is **only used in case of a failure in the primary path.**” *Id.* at pg. 152 (emphasis in original).

133. On July 30, 2010, the Examiner issued another Final Office Action, again rejecting the claims under § 103, this time over Chen, “The LSP Protection/Restoration Mechanism in GMPLS,” (“Chen”) in view of Voit and Blanchet. *Id.* at pgs. 102-111. The Examiner found that Chen teaches each element of the Challenged Claims other than a “configuration acknowledgement,” and that Chen’s teachings were in the context of a GMPLS environment, as opposed to a Pseudowire environment. *Id.* at pgs. 104-106. The Examiner further found, however, that Voit teaches Pseudowire and Pseudowire protection; and that Blanchet teaches the use of a “configuration acknowledgement.” The Examiner further found that it would have been obvious to a person of skill in the art to combine Chen with Voit because MPLS and Pseudowire are both point-to-point virtual links, because they are in the same field of endeavor (network transfer), and because they are directed to the same problem (data traffic protection) within that field. *Id.* at pgs. 105-106. The Examiner further found that it would have been obvious to combine Chen and Voit with Blanchet to send an “ACK message” indicating the acceptance of the configuration parameters to make the system more reliable, as all of the references are in the field of network transfer. *Id.* at pg. 106.

134. On September 27, 2010, the applicant submitted Remarks in response to the last Final Office Action. *Id.* at 096-099. The applicant did not dispute the Examiner’s finding that Voit disclosed Pseudowires and Pseudowire protection, that Blanchet disclosed receiving a configuration acknowledgement, nor that it would have been obvious to one of skill in the art to combine Chen with Voit and Blanchet. Instead, the applicant argued that Chen does not disclose the element of “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” *Id.*

135. In support of this argument, the applicant asserted that “the backup LSP is idle. Since the backup LSP is idle, **no traffic can exist on the backup LSP.**” *Id.* at 097 (emphasis added). Directly contradicting himself, the applicant went on to assert that “the resources allocated to the backup LSP may be used by other LSPs until the primary fails.” *Id.* at 098.

136. In my opinion, the applicant’s argument is technically inaccurate because the backup LSP clearly cannot be idle if it is being used by other LSPs until the primary LSP fails. The resources of a link are part of the link itself. Thus, in the context of the GMPLS system described by Chen, there is no other meaning to a disclosure that the resources allocated to the backup LSP may be used by other LSPs until the primary fails than that the backup path is used to transmit lower-priority traffic during normal operation. In other words, as a practical matter, stating that the resources are used for other LSPs means that the link is sending other LSP traffic while waiting for the primary link to fail. From a traffic engineering perspective, this situation is desirable and increases efficient use of resources.

137. Moreover, the protocols used to configure the portion of Chen cited by the Examiner (i.e., RSVP-TE) specifies exactly the opposite of what the applicant argued.

138. On October 14, 2010, the Examiner issued an Advisory Action rejecting the applicant's arguments. In particular, the Examiner noted that the backup path disclosed in Chen is not idle because Chen discloses that **“resource allocated for a backup LSP may be used by an LSP that has lower priority until the primary LSP fails.”** *Id.* at 095 (emphasis added).

139. On October 21, 2010, the applicant submitted a Pre-Appeal Brief Request for Review, and again argued that Chen fails to disclose “preempt[ing] **existing** traffic on the standby Pseudowire” based **“on the priority for the standby Pseudowire.”** *Id.* at 080 (emphasis added). The applicant argued that Chen does not disclose preempting “existing traffic” because either (1) the resources for the backup link are allocated for use by other LSPs with lower priorities, or (2) the backup path is dedicated to the primary path in a 1+1 scheme. *Id.* at 082-083. **The applicant also reiterated that Chen teaches preempting the use of prioritized resources, not preempting existing traffic.** *Id.* at 083.

140. On November 17, 2010, however, the Examiner issued a Notice of Panel Decision from Pre-Appeal Brief Review, stating that a conference had been held and that the rejection is withdrawn. *Id.* at 079-080. A Notice of Allowance was mailed on December 2, 2010, stating that the application was being allowed because the prior art did not teach “accepting the Pseudowire protection configuration parameter by the destination node, using the standby Pseudowire that is configured based at least in part on the Pseudowire configuration parameter, and determining whether to **preempt existing traffic** on the standby Pseudowire, wherein the **determination is based at least in part on the priority for the standby Pseudowire.**” *Id.* at 045-050 (emphasis added).

141. It should be noted that the applicant's arguments that the "1+1 protection scheme" of Chen does not disclose "existing traffic on the standby" is directly contrary to the Patent Owner's current construction of "existing traffic on the standby Pseudowire." Indeed, based on my review of the Patent Owner's infringement contentions, it is clear that the Patent Owner has done an about face from the positions it took regarding Chen and is now pointing to the duplicative traffic in a "1+1 protection scheme" as comprising the "existing traffic on the standby." *See* App. 25 (BSL's Preliminary Infringement Contentions) at '652 Chart, page 8.

142. In any event, Chen elsewhere discloses a variety of other protection schemes (e.g., 1+1, 1:1, 1:n, and M:N) that rebut the applicant's argument that the backup path is idle (either because it is not carrying any traffic or because it is carrying duplicative traffic, which the applicant does not view as "existing traffic"). *See, e.g.*, App. 7 (Chen) at pgs. 17, 53, and 54.

143. However, even if the applicant's arguments regarding Chen were correct, I have found multiple other pieces of prior art that do disclose the "preemp[tion] of existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire." A list of these references is attached as Appendix 8. *See also* App. 5 (RFC 3386), App. 15 (Owens) and App. 6 (Halabi) disclose this element.



## VI. CLAIM CONSTRUCTION

144. I understand that, for purposes of the accompanying petition for *Inter Partes* Review of the '652 patent (“Petition”), the Challenged Claims must be given their broadest reasonable interpretations in light of the specification of the '652 patent.

145. I understand that the parties have agreed to the following constructions in the Concurrent Litigation, which are relevant to each of the Challenged Claims:

<b>Pseudowire protection configuration parameter</b>	data structure with one or more fields that specify certain protection properties associated with a Pseudowire
<b>protection property</b>	field of data that corresponds to a protection scheme, protection type, domain type, and/or priority

App. 26 (Joint Claim Construction Statement) at 2.

146. The Challenged Claims each include the element of a “**standby Pseudowire**.” I understand that, in the Concurrent Litigation, BSL has not proposed a construction for this term, but has proposed constructions for a number of terms that include within them the term “standby Pseudowire.” In each of these constructions, BSL takes the position that a “standby Pseudowire” should be construed as an “emulation of a native service over a network that is used in the event of a network failure.” *Id.* at Ex. A, page 3. Based on my review of BSL’s infringement contentions, it is apparent that BSL is contending that Virtual Private LAN Services that traverse a service provider’s network over an MPLS LSP is included within the scope of “pseudowires.” App. 25 (BSL’s Preliminary Infringement Contentions) at page 2. Thus, the broadest reasonable interpretation

of “standby Pseudowire” should include at least “an emulation of a native service over a network that is used in the event of a network failure,” including, e.g., “VPLS” that traverse an MPLS LSP.

147. The Challenged Claims each require the standby Pseudowire to have a “**priority**” that is used, at least in part, to determine whether to preempt existing traffic on the standby Pseudowire. I understand that, in the Concurrent Litigation, BSL contends that “priority” means “preference.” In addition, I have reviewed BSL’s Preliminary Infringement Contentions. It is clear from those contentions that BSL is interpreting “priority” to encompass the mere designation of a Pseudowire as either a “primary” or “backup” Pseudowire, and is contending that “priority” does not require any separate “preference” designation. *See, e.g.*, App. 26 (Joint Claim Construction Chart) at Ex. A, at pg. 1; App. 25 (BSL’s Preliminary Infringement Contentions) at ’652 chart, page 9 (“the determination is based on the status (*e.g.*, primary vs. standby) of the two pseudowires. The primary and standby status of the two pseudowires indicate the respective priorities . . . with primary being higher than standby.”). Thus, the broadest reasonable interpretation of “priority” should include at least a “preference,” including, e.g., the designation of a Pseudowire as either “primary” or “backup” (or some equivalent, such as “standby” or “secondary”).

148. I further understand that Juniper has proposed that “priority” instead be construed as “a preference level for determining whether traffic on a Pseudowire should be preempted during a network failure that is different from its designation as a primary or standby Pseudowire” because the applicant repeatedly distinguished prior art that disclosed the designation of a path as primary versus standby as failing to disclose a “priority.” App. 3 (’652 File History) at 097-098 (distinguishing Chen’s primary/backup scheme for LSPs as disclosing “prioritized resources,” not a “priority” for the path); 152 (distinguishing Saleh’s

primary/secondary path scheme because it fails to disclose a “priority” that is assigned to the secondary path). As shown below, it is my opinion that the Challenged Claims are unpatentable under either Juniper’s construction or BSL’s construction.

149. The Challenged Claims each require receiving a “**Pseudowire configuration acknowledgment**” that indicates whether the Pseudowire configuration parameters have been accepted by the destination node. I understand that, in the Concurrent Litigation, BSL has proposed that this term means “an indication of whether the destination node accepts the standby Pseudowire,” and has contended that the mere execution and completion of a configuration command on a source node is sufficient to satisfy the this claim limitation. *See, e.g.*, App. 26 (Joint Claim Construction Chart) at Ex. A, at pg. 2; App. 25 (BSL’s Preliminary Infringement Contentions) at 7. Thus, the broadest reasonable interpretation of “Pseudowire configuration acknowledgement” should include at least “an indication of whether the destination node accepts the standby Pseudowire,” including, e.g., the execution and completion of a configuration command on a source node.

150. I further understand that Juniper has proposed that “Pseudowire configuration acknowledgment” should instead be construed as “a message from the destination node that is sent in response to the Pseudowire configuration parameter message.” As shown below, it is my opinion that the Challenged Claims are unpatentable under either Juniper’s construction or BSL’s construction.

151. The Challenged Claims each require “**determining whether to preempt existing traffic on the standby Pseudowire . . .**” I understand that, in the Concurrent Litigation, BSL has proposed that the term “existing traffic on the standby Pseudowire” should mean “working traffic transmitted on the [standby Pseudowire],” and has interpreted “working traffic” to include traffic that is

duplicative of the traffic that is being sent on a primary Pseudowire in a 1+1 scheme. *See, e.g.*, App. 26 (Joint Claim Construction Chart) at Ex. A, at pg. 3; App. 25 (BSL’s Preliminary Infringement Contentions) at 8. I further understand that BSL has interpreted “determining whether to preempt existing traffic” to encompass the act of dropping traffic from the standby Pseudowire during normal operation. *Id.* (“When the local PE router accepts traffic from the primary pseudowire and drops traffic from the standby pseudowire . . . , traffic from the primary pseudowire preempts traffic from the standby pseudowire.”).

152. I also understand that Juniper has proposed that “existing traffic on the standby Pseudowire” should instead be construed as “network traffic that is carried by the standby Pseudowire during normal operation that is different from the traffic that is carried on the protected Pseudowire” because the applicant repeatedly distinguished prior art that disclosed a 1+1 protection scheme. App. 3 (’652 File History) at 097-098 (distinguishing Chen’s 1+1 protection); 152 (distinguishing Saleh’s 1+1 protection). As shown below, it is my opinion that the Challenged Claims are unpatentable under either Juniper’s construction or BSL’s construction.

153. I understand that these statements by BSL may be used by the Patent Office to interpret claim language at issue. As discussed below, the Challenged Claims are certainly anticipated or obvious under BSL’s claim interpretation. The challenged claims are also anticipated or obvious under a narrower interpretation.

## VII. ANTICIPATION AND OBVIOUSNESS BASED ON PRIOR ART

### A. PRIOR ART

154. As discussed above, there are numerous prior art references that disclose the elements of the '652 patent. For purposes of my detailed analysis, I have identified eight key prior art references. The key references include:

155. **U.S. Patent Pub. No. 2004/0156313** to Hofmeister et al. (“Hofmeister”), which was filed on February 3, 2004, published on August 12, 2004 and issued as U.S. Patent No. 7,417,950 B2 on August 26, 2008. I note that the inventor of the '652 patent, Ping Pan, is a co-inventor on Hofmeister, but that Hofmeister is not a named inventor on the '652 patent. As such, Hofmeister is “by another.” Hofmeister is prior art under at least §§ 102(a) and 102(e).

156. **Request for Comments 3386** (“RFC 3386”) was published by the IETF and publically available no later than November 2002. It is prior art under at least 35 U.S.C. §§ 102(a) and (b).

157. **U.S. Patent No. 7,804,767 B1** to Owens et al. (“Owens”) was filed on October 25, 2000 and issued on September 28, 2010. App. 15 (Owens). Owens is prior art to the '652 patent under at least § 102(e).

158. **Request for Comments 3209** (“RFC 3209”) was published by the IETF and publically available no later than December 2001. It is prior art under at least 35 U.S.C. §§ 102(a) and (b).

159. “**Metro Ethernet**” is a book by Sam Halabi that was published by Cisco Press on October 1, 2003 (“Halabi”). It was copyrighted in 2003 and is catalogued in the Library of Congress under number 2002103527. *See, e.g.*, App. 27 (BSL RFA Responses) at 1 (admitting that Halabi is prior art). It is prior art under at least 35 U.S.C. §§ 102(a) and (b).

160. “**The LSP Protection/Restoration Mechanism in GMPLS**” is an article by Ziyang Chen (“Chen”) that is dated October 1, 2002 and was publicly

available no later than 2003. *See, e.g.*, App. 27 (BSL RFA Responses) at 1 (admitting that Chen is prior art). Chen was discussed in the file history of the '652 patent. Chen is prior art under at least 35 U.S.C. §§ 102(a) and (b).

161. **U.S. Patent No. 7,305,481 B2** to Blanchet et al. (“Blanchet”) was filed on January 3, 2007, published on July 8, 2004, and issued on December 4, 2007. Blanchet was discussed in the file history of the '652 patent. It is prior art under at least §§ 102(a) and (e).

162. **U.S. Patent Pub. No. 2006/0047851 A1** to Voit et al. (“Voit”) was filed on August 25, 2004 and issued as U.S. Patent No. 7,643,409 B2 on January 10, 2010. Voit was discussed in the file history of the '652 patent. It is prior art under at least 35 U.S.C 102(e)

163. As discussed below, it is my opinion that: (1) claims 1, 9 and 14 are anticipated by Hofmeister; (2) the Challenged Claims are obvious over Hofmeister in view of RFC 3386 and Owens; (3) the Challenged Claims are anticipated by RFC 3386; (4) the Challenged Claims are obvious over RFC 3386 in view of RFC 3209; (5) the Challenged Claims are anticipated by Halabi; (6) the Challenged Claims are obvious over Halabi alone; (7) the Challenged Claims are obvious over Halabi in view of RFC 3386 and Owens; and (8) the Challenged Claims are obvious over Chen in view of Voit and Blanchet.

**B. Hofmeister anticipates Claims 1, 9 and 14 under 35 U.S.C. §§ 102(a) and 102 (e)**

164. U.S. Patent No. 7,417,950 B2 to Hofmeister et al. (“Hofmeister”) was filed on February 3, 2004, published on August 12, 2004 and issued on August 26, 2008.

165. Hofmeister is entitled “Method and Apparatus for Performing Data Flow Ingress/Egress Admission Control In A Provider Network” and discloses a technique for setting up and managing Pseudowires on a SONET network backbone.

166. Hofmeister teaches that “[d]etailed network resource information particular to each of the data flows is exchanged between provider edge nodes during the creation of pseudo-wires” and that “[b]y applying pseudo-wire shuffling and preemption techniques, the providers can make better use of their network resources.” App. 4 (Hofmeister) at Abstract.

167. It is my understanding that Hofmeister was submitted to the PTO during prosecution, but that it was submitted in an Information Disclosure Statement that cited to 60 other references and that was filed after the PTO had already issued a Notice of Allowance. Based on my review of the file history, it does not appear that the Examiner engaged in a substantive analysis of Hofmeister.

*1. Claim 1: A method of providing protection to network traffic, comprising,*

168. Claim 1 recites: “A method of providing protection to network traffic, comprising. . . .” Under BSL’s apparent claim construction, Hofmeister discloses this element.

169. For example, Figure 34 of Hofmeister contemplates a “Backup” mechanism for “PW1 traffic”:

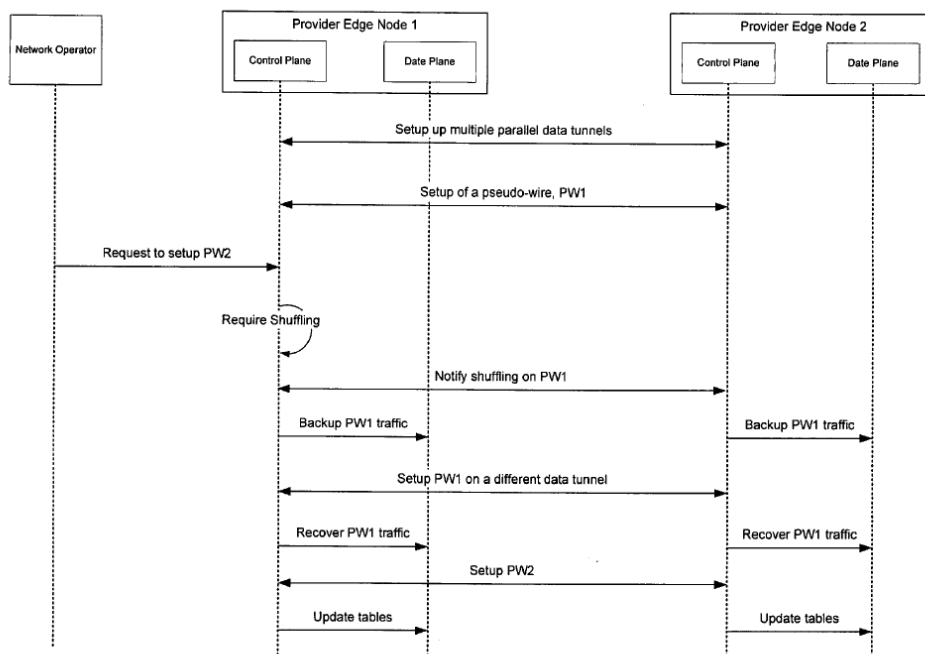


Figure 34

170. As another example, Hofmeister notes that building pseudo-wires over SONET is desirable because of the “rich set of features for . . . traffic *restoration*, and link *protection*.” App. 4 (Hofmeister) at [0257]. As a further example, Hofmeister teaches that, in some scenarios it is desirable “direct existing traffic to a backup link.” *Id.* at [0397].

- a. *sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire,*

171. Claim 1 further recites: “sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source



node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

172. For example, Hofmeister depicts the process of exchanging Pseudowire data service guarantee information between Provider Edge Nodes in Figure 23(b), which is an example of sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node (PE Node C) and a destination node (PE Node H):

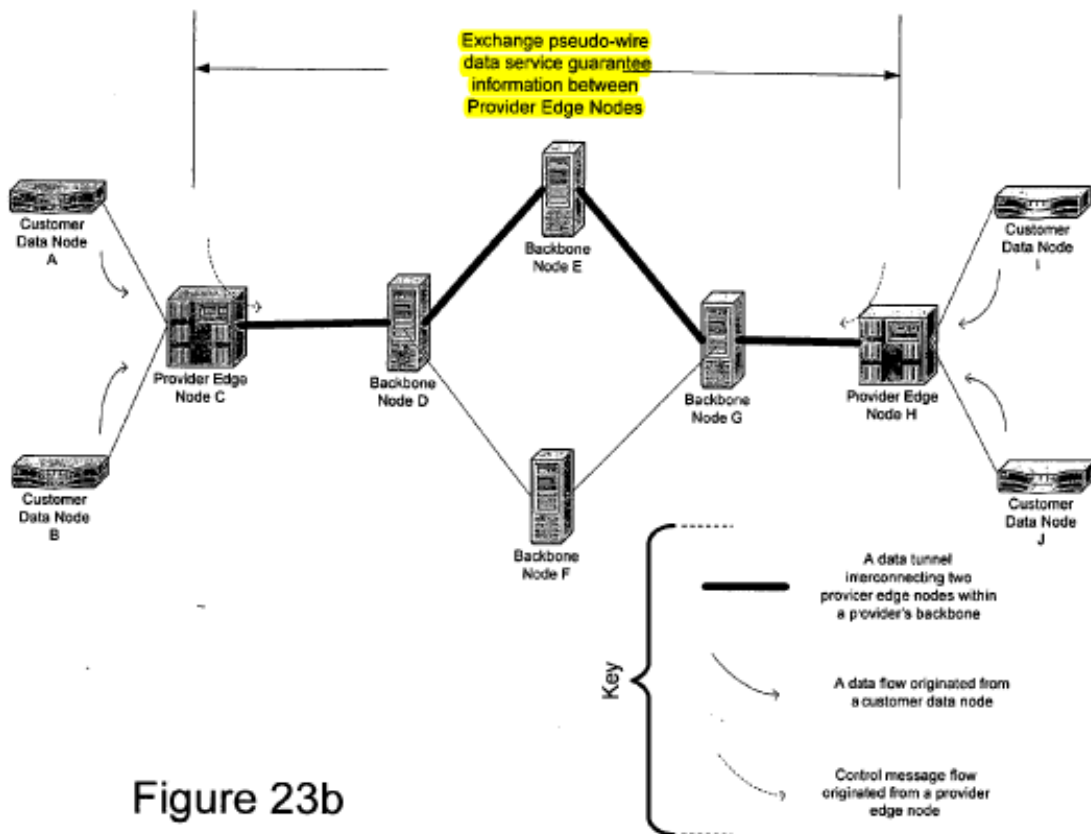


Figure 23b

App. 4 (Hofmeister) at Fig. 23(b); *see also id.* at [0046] (“**FIG. 23b** is another high-level network diagram illustrating the operation of service negotiation between Provider Edge Nodes according to the concepts of the invention”).

173. As another example, Hofmeister teaches that “control messages are used to discover and establish pseudo-wires” (*see* App. 4 (Hofmeister) at [0081]) and that two edge nodes can establish a Pseudowire by exchanging control messages. *See id.* at [0203] (“Node-1 and Node-2 will establish a peering session . . . .The method for session establishment is to inject control messages into the connection.”). This is another example of sending a Pseudowire protection configuration parameter (*i.e.*, “control message”) between a source node (*i.e.*, “Node-1”) and a destination node (*i.e.*, “Node-2”); *see also id.* at [0202] – [0211].

174. In a more detailed example, Hofmeister explains that setting up a Pseudowire involves the following steps:

- An operational optical connection between two provider edge nodes exists (*id.* at [0202]);
- Node-1 and Node-2 establish a peering session by injecting “control messages into the connection” that are encapsulated with an identifiable label (*id.* at [0203]);
- Once the peering session is established, the Network operator issues a data flow setup request to both nodes that include information about the parameters of the path, such as the relevant data interfaces, QoS (bandwidth) requirements (*id.* at [0204]);
- Node-1 and Node-2 exchange “control messages” and negotiate the labels to be used by the data flows (*id.* at [0210]);
- Upon completion of label negotiation, Node-1 and Node-2 update the data-plane with the label information by populating their respective packet filter tables and circuit filter tables (*id.* at [0211]).

175. As another example, in Figure 34, Hofmeister shows that Node 1 and Node 2 “setup **multiple** parallel data tunnels” and depicts that “Backup PW1 traffic” can be done by “Setup PW1 on a different data tunnel.” *Id.* at Fig. 34. This is an example of “configuring a standby Pseudowire.”

176. In another embodiment, Hofmeister explains that the “pseudo-wire information” that is exchanged between the provider edge nodes can include additional parameters, such as Committed Information Rate (CIR), Class, Setup Priority and Holding Priority. These parameters are assigned to each flow, as shown in Figure 27:

280

CIRCUIT FILTER TABLE					
CIRCUIT FILTER (DATA TUNNEL, LABEL)	OUTGOING DATA INTERFACE	CIR	CLASS	SETUP PRIORITY	HOLDING PRIORITY
CIRCUIT FILTER-1 (SONET VCG 3, LABEL 20000)	DATA PORT 1	50 Mb/Sec	AF-1	3	3
CIRCUIT FILTER-2 (SONET VCG 3, LABEL 20001)	DATA PORT 2	8 Mb/Sec	EF	3	3
CIRCUIT FILTER-3 (OPTICAL INTERFACE 1, LABEL 300)	DATA PORT 10	10 Gb/Sec	AF-3	3	3
CIRCUIT FILTER-4 (SONET VCG 5, LABEL 12000)	DATA PORT 1	100 Mb/Sec	AF-1	1	1
CIRCUIT FILTER-5 (SONET VCG 3, LABEL 3)	DATA PORT 2	1 Gb/Sec	AF-1	5	5

Figure 27

See App. 4 (Hofmeister) at Fig. 27; see also *id.* at Fig. 28, [0291], [0292], [0293], [0294] and [0295]; see also [0381]. These parameters are examples of “Pseudowire protection configuration parameters.”

177. As another example, Hofmeister discloses that the parameters can include information about the domain. *See, e.g., id.* at [0222] and [0261].

178. It is my understanding that the parties have agreed in the Concurrent Litigation that the term “Pseudowire protection configuration parameter” should mean “data structure with one or more fields that specify certain protection properties associated with a Pseudowire.” Under this proposed construction, the “control messages” that are used to negotiate the Pseudowire labels and exchange information about the data flow parameters are “Pseudowire protection configuration parameters.” And, because these messages are sent between Node-1 and Node-2, they disclose “sending a Pseudowire protection configuration parameter . . . between a source node and a destination node.”

179. It is my understanding that the parties have agreed in the Concurrent Litigation that the term “protection property” should be construed as “field of data that corresponds to a protection scheme, protection type, domain type, and/or priority.” Under this construction, the “Setup Priority” and “Holding Priority” disclosed by Hofmeister comprise “a protection property” because they comprise a “field of data” that corresponds to a “priority” for the Pseudowire that is being set up.

180. Moreover, the “Setup Priority” and “Holding Priority” disclosed in Hofmeister appear to be the same as the examples of a “priority” that are disclosed in the ’652 patent. *See App. 2* (’652 patent) at 6:57 – 7:5 (discussing holding and setup priorities). As such, they would certainly be included within the broadest reasonable interpretation of a “protection property.”

181. In one example, the process for configuring a Pseudowire that is disclosed by Hofmeister could be used to set up a standby or backup Pseudowire. Indeed, Hofmeister states that, in at least some circumstances, it would be desirable to be able to switch a data flow to a backup link. *See, e.g., id.* at [0397], [0399]

(noting desirability of having a “backup” link); [0257] (“[i]n many cases, SONET connections are well provisioned with a rich set of features for network resource allocation, traffic restoration, and link protection, etc.”).

*b. the protection property including a priority for the standby Pseudowire*

182. Claim 1 further recites: “the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

183. For example, Hofmeister teaches that control messages used to set up a Pseudowire can include a “Setup Priority” (App. 4 (Hofmeister) at [0407]) and/or a “Holding Priority” (*id.* at [0408]).

184. Figure 27 and Figure 28 of Hofmeister also show how a “Setup Priority” and “Holding Priority” can be assigned to each Pseudowire (flow) and how those priorities can be used during operation. *Id.* at Fig. 27 and Fig. 28.

185. It is my understanding that BSL has proposed that “priority” be construed to simply mean “preference.” Under this broad construction, the “Setup Priority” and “Holding Priority” disclosed by BSL would comprise a “priority.”

186. Moreover, I note that the “Setup Priority” and “Holding Priority” disclosed in Hofmeister appear to be the same as the examples of a “priority” that are disclosed in the ’652 patent. *See* App. 2 (’652 patent) at 6:57 – 7:5 (discussing holding and setup priorities). As such, they would certainly be included within the broadest reasonable interpretation of a “priority.”

- c. *receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

187. Claim 1 further recites: “receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

188. For example, Hofmeister shows that the source node and the destination engage in two-way communication regarding the Pseudowire setup in Figure 23b, as it depicts “Exchange pseudo-wire data service guarantee information between Provider Edge Nodes”:

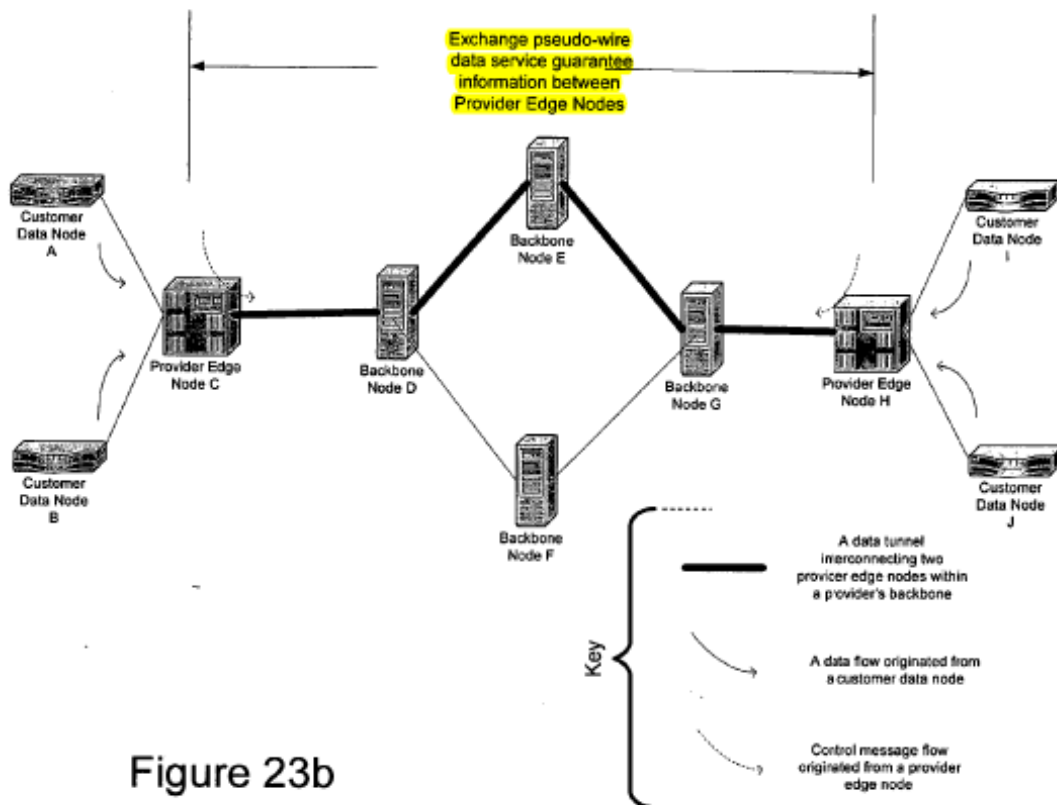


Figure 23b

189. As another example, Hofmeister teaches that “[a]s a part of pseudo-wire process as defined in LDP and draft-martini, the Terminating PE node will *acknowledge* the establishment of the new flow to the Initiating PE node.” App. 4 (Hofmeister) at [0372]; *see also id.* at [0384]. In this scenario, the “Terminating PE node” is the “destination node” and the “Initiating PE node” is the “source node.”

190. In addition to the explicit disclosure of an “acknowledgement,” Hofmeister also teaches that Pseudowires can be set up using LDP, RSVP-TE, and/or draft-martini standards. *See, e.g.,* App. 4 (Hofmeister) at [0091], [0290]. At the time of the Hofmeister invention, one skilled in the art would understand that these specific protocols use a “configuration acknowledgement” that is sent by the destination node to the source node to indicate whether the parameters requested by the source node have been accepted under BSL’s apparent construction of this claim element.

191. For example, LDP is described in RFC 3036, which was published on the IETF website no later than January 2001. As described in RFC 3036, a Label-Switched Router (“LSR”) sends a Label Request to another LSR to initiate an LDP tunnel set-up. App. 10 (RFC 3036) at § 3.5.8. Further, “[t]he response by an LSR to receipt of a FEC label request from an LDP peer may involve one or more of the following actions: Transmission of a notification message to the requesting LSR indicating why a label mapping for the FEC cannot be provided; Transmission of a FEC label mapping to the requested LSR; Transmission of a FEC label request to the FEC next hop; Installation of labels for forwarding/switching use by the LSR.” *Id.* at A.1.1.1. Under the broadest reasonable interpretation of this claim element, each of these are examples of “configuration acknowledgements” that indicate whether the request parameters have been accepted or not. Thus, the use of “LDP”

protocol necessarily involves the use of a “configuration acknowledgement” indicating whether the requested parameters have been accepted.

192. As another example, the RSVP-TE protocol is described in RFC 3209, which was published by the IETF in December of 2001. In this protocol, a “sender node” sends a “Path Message” with a “LABEL\_REQUEST object” to a “destination node.” The Path Message contains information about the requested parameters for the LSP that is being set up. RFC 3209 explains that “[t]he destination node of a label-switched path response to a LABEL\_REQUEST by including a LABEL object in its response RSVP Resv message. The LABEL object is inserted in the filter spec list immediately following the filter spec to which it pertains.” App. 9 (RFC 3209) at § 2.2. Thus, the use of “RSVP-TE” protocol necessarily involves the use of a “configuration acknowledgement” indicating whether the requested parameters have been accepted.

193. As another example, the “draft-martini” standard for configuring a Pseudowire is described in “Pseudowire Setup and Maintenance using LDP” (v. 5) by Luca Martini, Eric Rosen, Nasser El-Aawar and Toby Smith, which was published in December 2003. App. 30 (Martini Draft). In § 5.2.3, that document explains:

“To begin the signaling procedure, a PE (PE1) [*i.e.*, the source node] that has knowledge of the other endpoint (PE2) [*i.e.*, the destination node] initiates the setup of the LSP in the incoming (PE2→PE1) direction by sending a Label Mapping message containing the FEC type 129 [*i.e.*, the configuration parameter]. The FEC element includes the SAII, AGI and TAI. . . . PE 2 interprets the message as a request to set up a Pseudowire whose endpoint (at PE 2) is the Forwarder identified by the TAI. . . . If PE2 cannot map the TAI to one of its Forwarders, then PE2 send a Label Release message to PE1,



with a Status Code meaning ‘invalid TAI,’ and the processing of the Mapping message is complete. . . If the Label Mapping Message has a valid TAI, **PE2 must decide whether to accept it or not.** . . . If PE2 decides to accept the Label Mapping message, then it has to make sure that an LSP is set up in the opposite (PE1 → PE2) direction. If it has already signaled for the corresponding LSP in that direction, nothing more need be done. Otherwise, **it must initiate such signaling by sending a Label Mapping message to PE1.**”

*Id.* at 5.2.3 (emphasis added).

194. The Label Mapping message that PE2 sends back to PE1 in response to the Label Mapping message from PE1 is a “configuration acknowledgement” message that indicates whether the requested parameters have been accepted by the destination node (*i.e.*, PE2) under the broadest reasonable interpretation of this claim element. Thus, the use of “draft-martini” to set up a Pseudowire necessarily involves the use of a “configuration acknowledgement” indicating whether the requested parameters have been accepted.

195. These references concerning RSVP-TE, LDP, and draft-martini were standard background knowledge for those of ordinary skill in the art and are being cited as other written evidence solely to help explain Hofmeister.

196. Thus, Hofmeister’s reference to these signaling protocols discloses “receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted,” under BSL’s apparent claim constructions.

d. accepting the Pseudowire protection configuration parameter by the destination node;

197. Claim 1 further recites: “accepting the Pseudowire protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

198. For example, in Figures 23b and 34, Hofmeister shows the process of setting up a Pseudowire according to specific protection configuration parameters, which involves the sending, receiving and accepting of those parameters. See App. 4 (Hofmeister) at Figs. 23b and 34.

199. As another example, Figure 27 shows a list of circuits and protection parameters for each Pseudowire (data flow) that is set up:

280

CIRCUIT FILTER TABLE					
CIRCUIT FILTER (DATA TUNNEL, LABEL)	OUTGOING DATA INTERFACE	CIR	CLASS	SETUP PRIORITY	HOLDING PRIORITY
CIRCUIT FILTER-1 (SONET VCG 3, LABEL 20000)	DATA PORT 1	50 Mb/Sec	AF-1	3	3
CIRCUIT FILTER-2 (SONET VCG 3, LABEL 20001)	DATA PORT 2	8 Mb/Sec	EF	3	3
CIRCUIT FILTER-3 (OPTICAL INTERFACE 1, LABEL 300)	DATA PORT 10	10 Gb/Sec	AF-3	3	3
CIRCUIT FILTER-4 (SONET VCG 5, LABEL 12000)	DATA PORT 1	100 Mb/Sec	AF-1	1	1
CIRCUIT FILTER-5 (SONET VCG 3, LABEL 3)	DATA PORT 2	1 Gb/Sec	AF-1	5	5

Figure 27

And Figure 28 also shows that the protection parameters are tracked in the Session Table for each Pseudowire (data flow):

SESSION TABLE				
SESSION (CONTROL MESSAGE ID)	OUTGOING DATA TUNNEL	ENCAPSULATION LABEL	CIR	CLASS
SESSION 1 (TCP SRC PORT 1345)	SONET VCG NUMBER 3	MPLS LABEL 3	1 Mb/Sec	EF
SESSION 2 (TCP SRC PORT 3456)	MPLS LSP 8	MPLS LABEL 3	2 Mb/Sec	EF
SESSION 3 (TCP SRC PORT 1998)	OPTICAL INTERFACE 1	MPLS LABEL 10000	N/A	EF

Figure 28

The inclusion of the protection parameters in the session table would necessarily require that those protection configuration parameters were accepted during the configuration process.

200. As another example, Hofmeister teaches that the control module of the terminating PE (*i.e.*, “destination node”) captures and forwards control messages for a new Pseudowire. App. 4 (Hofmeister) at [0369]. As discussed above, these “control messages” can contain data flow information, such as bandwidth, class and priority, for the Pseudowire that is being set up. Hofmeister further discloses that the local control module on the terminating PE assesses whether the terminating PE has the required bandwidth, and if the requested data flow CIR does not exceed the data port capacity, the terminating PE accepts the new Pseudowire. *Id.* at [0370]; [0384]. In this example, if the terminating PE accepts the new Pseudowire, it necessarily accepts the corresponding protection configuration parameter.

- e. using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

201. Claim 1 further recites: “using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration

parameter.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

202. For example, Hofmeister discloses the use of a Pseudowire that has been configured with a Setup Priority and Holding Priority. App. 4 (Hofmeister) at [0294], [0295].

203. By way of example, Hofmeister explains that the “Setup Priority . . . describes the priority of a given pseudo-wire with respect to taking resources. This value is used in deciding whether this pseudo-wire can preempt another pseudo-wire . . . . During pseudo-wire provision, when th[ere] is no sufficient amount of network resource, a data flow with higher Setup Priority value can preempt the pseudo-wires with lower priority from a data tunnel” See App. 4 (Hofmeister) at [0294]. Further to this example, Hofmeister explains that “Holding Priority . . . describes the priority of the pseudo-wire with respect to holding resources. The Holding Priority is used in deciding whether this pseudo-wire can be preempted by another pseudo-wire.” *Id.* at [0295].

204. Further to this example, Hofmeister provides a detailed explanation regarding how a Pseudowire can be used in operation based on the Setup and Holding Priorities that were assigned during configuration. See App. 4 (Hofmeister) at [0406], [0407], [0408] [0411], [0412], [0413]:

[0406] In the invention, two basic priority classes may be used for each flow: Setup Priority and Holding Priority which are defined above in detail and further discussed below.

[0407] Setup Priority is the relative importance (ranking) of a new pseudo-wire with respect to taking resources from other pre-established pseudo-wires.

[0408] Holding Priority is the relative importance (ranking) of an existing pseudo-wire with respect to holding the resources from being taken away or pre-empted by another pseudo-wire requesting admission.

[0411] The preemption steps illustrated in **FIG. 30** and discussed above may utilize the more detailed preemption algorithm illustrated in **FIG. 33**. The preemption algorithm requires an input (900) providing the information on a pseudo-wire's resource requirement, set-up priority and the data tunnel or interface where preemption will take place. Note that this algorithm is applied at both ingress and egress interfaces. At ingress, preemption may take place in a data tunnel, whereas preemption can remove less important flows from a data interface at an egress point.

[0412] The algorithm begins in earnest by searching (905) for all the flows having a holding priority less than the setup priority of the new flow. If at the ingress point, the searching (905) is done within the packet filter table 260.

[0413] With this information in hand, the method may then determine (910) if the combined resources from the selected flow(s) having a holding priority less than the setup priority are not enough to accommodate the new flow. If so, the algorithm will fail (915) the preemption process because the new flow simply cannot be accommodated according to the relative priority levels and resource demands. Otherwise, the algorithm will select (920) a set of flows that can accommodate the new flow and which may be preempted according to the relative setup and holding priorities. The combined resources from the selected (920) flows will be larger or equal to the resources required by the new flow. The actual selection mechanism (920) may be based on provider policy. For example, only the smallest flows would be preempted, or the flow selection can be random.

205. Thus, as an example, if Pseudowire A has been assigned a Set-Up Priority of “6” via the exchange of control messages during configuration, and Pseudowire B has been set up as a “standby” or “backup” Pseudowire with a Holding Priority of “2” via the exchange of control messages during configuration,

Hofmeister teaches that, if there were not enough resources to set up Pseudowire A, then the system would preempt Pseudowire B because Pseudowire B has a lower priority. This is an example of “using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter.”

206. As another example, Figure 34 demonstrates the use of “shuffling” in connection with the use of PW1. The concept of “shuffling” is based on the various protection configuration parameters assigned to PW1 during the configuration process, such as Setup Priority and Holding Priority.

*f. determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

207. Claim 1 further recites: “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

208. For example, Hofmeister teaches the use of a preemption algorithm to determine whether to preempt the traffic flowing on a Pseudowire in order to accommodate a request for a new Pseudowire to be set up. App. 4 (Hofmeister) at [0410], [0411]. The algorithm uses the relative Holding Priority and Setup Priority that was assigned to each Pseudowire during configuration to make the determination as to whether to preempt the data flowing on the existing Pseudowire in order to accommodate the new Pseudowire. *Id.* at [0412], [0413].

209. It is my understanding that BSL has proposed that the phrase “existing traffic on the standby Pseudowire” be interpreted to mean “working traffic on the standby Pseudowire.” Under BSL’s proposed construction, the traffic flow on a

current Pseudowire that was set up as a “backup” or “standby” would constitute “existing traffic.” Moreover, the use of the preemption algorithm to determine whether to preempt a Pseudowire that has been set up as a standby or backup would constitute “determining whether to preempt existing traffic on the standby Pseudowire based, at least in part, on the priority for the standby Pseudowire.”

210. Thus, returning to the example I provided above, where Pseudowire A has been assigned a Setup Priority of “6” via the exchange of control messages during configuration, and Pseudowire B has been set up as a standby or backup with an assigned Holding Priority of “2” via the exchange of control messages during configuration, and where the system would preempt standby Pseudowire B if Pseudowire A makes a setup request and there is not enough bandwidth, the preemption of standby Pseudowire B would comprise “preempting existing traffic” and the preemption determination would be based on the associated priority that was assigned to the standby Pseudowire during configuration.

211. In sum, it is my opinion that Hofmeister discloses each and every element of claim 1 of the ’652 patent under BSL’s apparent claim constructions, and thus Hofmeister anticipates claim 1.

#### **Claims 9 and 14**

212. I note that the limitations of independent claim 9 and claim 14 are nearly identical to claim 1, except for the fact that claim 9 is a system claim and claim 14 is a computer program claim.

9. *Claim 9: A system for providing protection to network traffic, comprising a processor configured to:*

213. Claim 9 recites: “a system for providing protection to network traffic, comprising a processor.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

214. I incorporate by reference my comments from claim 1 above.

215. In addition, Hofmeister teaches that the techniques disclosed for configuring Pseudowires can be implemented on a processor. App. 4 (Hofmeister) at [0299] (control module can be implemented on a microprocessor, FPGA, or ASIC).

- a. *send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;*

216. Claim 9 further recites: “send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

217. I incorporate by reference my comments from claim element 1(a) and 1(b) above.

- b. *receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

218. Claim 9 further recites: “receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, Hofmeister discloses this element.



219. I incorporate by reference my comments from claim element 1(c) above.

*c. accept the Pseudowire protection configuration parameter by the destination node;*

220. Claim 9 further recites: “accept the Pseudowire protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

221. I incorporate by reference my comments from claim element 1(d) above.

*d. use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

222. Claim 9 further recites: “use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

223. I incorporate by reference my comments from claim element 1(e) above.

*e. determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

224. Claim 9 further recites: “determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

225. I incorporate by reference my comments from claim element 1(e) above.

226. In sum, it is my opinion that Hofmeister discloses each and every element of claim 9 of the '652 patent under BSL's apparent claim constructions, and thus Hofmeister anticipates claim 9.

*14. Claim 14: A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for:*

227. Claim 14 recites: "a computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions." Under BSL's apparent claim construction, Hofmeister discloses this element.

228. I incorporate by reference my comments from claim 1 above.

229. In addition, Hofmeister teaches that the techniques disclosed for configuring PWs can be implemented on a processor. App. 4 (Hofmeister) at [0299] (control module can be implemented on a microprocessor, FPGA, or ASIC). This suggests that the invention would comprise a computer program product that is embodied in a computer readable storage medium and that comprises computer instructions.

*a. sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby*

*Pseudowire, the protection property including a priority for the standby Pseudowire;*

230. Claim 14 further recites: “sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

231. I incorporate by reference my comments from claim elements 1(a) and 1(b) above.

*b. receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

232. Claim 14 further recites: “receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

233. I incorporate by reference my comments from claim element 1(c) above.

*c. accept the Pseudowire protection configuration parameter by the destination node;*

234. Claim 14 further recites: “accept the Pseudowire protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

235. I incorporate by reference my comments from claim element 1(d) above.

*d. using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

236. Claim 14 further recites: “using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

237. I incorporate by reference my comments from claim element 1(e) above.

*e. determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

238. Claim 14 further recites: “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Hofmeister discloses this element.

239. I incorporate by reference my comments from claim element 1(f) above.

240. In sum, it is my opinion that Hofmeister discloses each and every element of claim 14 of the ’652 patent under BSL’s apparent claim constructions, and thus Hofmeister anticipates claim 14.

241. Thus, it is further my opinion that Hofmeister anticipates claims 1, 9, and 14 of the ’652 patent.

**C. Hofmeister in view of RFC 3386 and Owens renders Claims 1-5, 8-11, 13-15 and 17 obvious under 35 U.S.C. § 103**

242. If certain aspects recited in claims 1, 9 and 14 are not deemed to be disclosed or inherent over Hofmeister alone, it is my opinion that the inclusion of those aspects certainly would be obvious over Hofmeister in view of RFC 3386 and Owens. The combination of Hofmeister with RFC 3386 and Owens also would render dependent claims 2-5, 8, 10-11, 13, 14, and 17 obvious.

243. As I described above, Hofmeister teaches a detailed method for signaling and managing PWs over a SONET backbone. App. 4 (Hofmeister) at Abstract, [0086]. Hofmeister expressly notes that the disclosed invention “leverages [] conventional technologies” drawn from IETF Internet Drafts and RFCs, including the Martini draft regarding PWs (App. 4 (Hofmeister) at [0010]), the Swallow draft regarding RSVP-TE for Fast-Reroute (*id.* at [0014]), and other technologies such as MPLS, OIF UNI, Virtual Concatenation, LCAS and GFP (*id.* at [0016]).

244. Hofmeister further states that it utilizes configuration parameters from conventional IETF industry standards, such as CIR (RFC 2697/ 2698), Traffic Class (RFC 2475 on Internet DiffServ), and Setup/Holding Priorities (multiple RFCs regarding RSVP-TE protocol for MPLS). *Id.* at [0296]; *see also* App. 11 (RFC 2697); App. 12 (RFC 2698); App. 13 (RFC 2475); App. 9 (RFC 3209); App. 14 (RFC 4090).

245. RFC 3386 was published by the IETF. It describes various configuration parameters that can be used to provide traffic protection in a wide range of networks, including SONET, MPLS, GMPLS, and Pseudowire environments.

246. The protection techniques described in RFC 3386 were widely-known and utilized by the industry at the time of the Hofmeister patent. As a specific

example of this, Owens (entitled “Protection/ Restoration of MPLS Networks”) is a patent that was filed by a group of engineers from Tellabs who were involved in the IETF. It provides a more detailed description of how the protection techniques in RFC 3386 can be applied to an MPLS environment. For example, Owens teaches that, in an MPLS network, “a working path carries data from a starting point or node to a destination point or node via a working path . . . . MPLS system reliability is enhanced by way of a protection path, over which data can be carried from the starting point to the destination point upon a detected failure along the working path. App. 15 (Owens) at Abstract.

247. Owens also describes various specific protection features in great detail. For example, Owens teaches that the protection path can be configured using “dynamic” protection (which is akin to the “cold” mode discussed in the ’652 patent) or “pre-negotiated” protection (which is akin to the “hot” or “warm” mode discussed in the ’652 patent). *Id.* at 5:1-29. Owens also discloses various “protection modes,” such as revertive or non-revertive, as well as a number of “protection switching options, such as “1+1 protection,” and “1:1, 1:n and n:m Protection.” *Id.* at 6:16 – 7:15.

248. Given the prominence of the protection methods discussed in RFC 3386 and Owens, and their close relationship to the protocols from which the Hofmeister invention explicitly derives, it would have been obvious to apply the protection techniques and parameters described in RFC 3386 and Owens to the specific Pseudowire environment described by Hofmeister.

249. In fact, combining Hofmeister with RFC 3386 and Owens would have required nothing more than applying the known protection techniques of priority and preemption to the known network environment of Pseudowires in the same manner that the technique is applied to other network environments and to achieve the same general result (traffic protection).

250. Moreover, a person of skill in the art would have been motivated to combine the references. Indeed, Hofmeister and RFC 3386 are related to the same very narrow field of hybrid L2-L3 internet protocols. As discussed above, the Pseudowire concept was initiated by the “Martini-draft” which was a draft submitted to the IETF by Luca Martini from Cisco Systems. Cisco wanted to solve the problem of providing L2 services over geographically dispersed areas by transporting L2 frames over MPLS, which is what prompted the “Martini-draft” and the standardization of a protocol for Pseudowires. A number of network vendors, including Nortel, Juniper, and Ciena were working along with service providers such as AT&T, Worldcom and Level-3 to offer this service to the customers, and they were therefore constantly collaborating in the IETF forum regarding the relevant standards and protocols. The process involved the individuals from these companies proposing many draft documents, presenting them to the other members of the IETF community and attempting to get a consensus on the right solution. Indeed, for interoperability, it is necessary to get an agreement on the standard to be used by all parties involved (*i.e.*, both the network product providers and the service providers).

251. The fact that RFC 3386 was drafted by members of this tight-knit IETF community from the service provider side (*i.e.*, W. Lai of AT&T and D. McDysan of Worldcom) and Hofmeister and Owens were drafted by members from the network product provider side (Tad Hofmeister and Ping Pan of CIENA Corp. and K. Owens, S. Makan, C. Huang, and V. Sharma of Tellabs) further evidences that one skilled in the art would have been motivated to combine the concepts discussed in RFC 3386 with the concepts discussed in Hofmeister.

252. A person of skill in the art would have also been motivated to combine Hofmeister and RFC 3386 because the references are directed to the same problem of obtaining greater control over configuration, management and

resiliency of network environments. Indeed, all three references are part of a relatively narrow body of literature on MPLS and Pseudowire set-up, traffic engineering and protection.

253. In addition, Hofmeister itself notes that a benefit of the claimed Pseudowire environment is that it can take advantage of the “rich set of features for network resource allocation, traffic *restoration*, and link *protection*” (App. 4 (Hofmeister) at [0257]) and that it allows “*protection* mechanisms” to be “triggered much faster thereby preventing data loss” (*id.* at [0137]). With respect to some of the disclosed embodiments, Hofmeister notes that it would be desirable to direct traffic to a **backup link**. *Id.* at [0397] (“direct existing traffic to a backup link (*e.g.* such as using protection bandwidth triggered via a conventional APS (automatic protection switch) protocol for SONET/SDH traffic”).

254. RFC 3386 and Owens would have been obvious places to look for the specific protection techniques and schemes that could be applied to achieve these protection benefits and goals. Indeed, RFC 3386 in particular was a prominent specification on protection methods.

255. Moreover, the '652 patent itself states that MPLS protection schemes (such as “MPLS Fast Reroute”) are relevant to Pseudowires and that they can be used to protect data being transmitted on a Pseudowire. App. 2 ('652 patent) at 1:49-64. As such, the '652 patent acknowledges the obviousness of and motivation to combine MPLS protection techniques with Pseudowires.

### **Claims 1, 9, 14**

256. If certain aspects recited in claims 1, 9 and 14 are not deemed to be disclosed or inherent over Hofmeister alone, it is my opinion that the inclusion of those aspects certainly would be obvious over Hofmeister in view of RFC 3386 and Owens.



**“standby Pseudowire”**

257. Claim 1, 9 and 14 recite in part: “configuring a standby Pseudowire.”

258. To the extent that Hofmeister is deemed to not expressly or inherently disclose this element, it is my opinion that this element would have been obvious in view of RFC 3386 and Owens.

259. For example, the use of a “standby” was a commonly-used concept and had been well-known for years. For example, RFC 3386 teaches the use of a “protection connection” to protect a “working” path. App. 5 (RFC 3386) at § 3.2.1.

260. Similarly, Owens describes a “Backup LSP (or Protection or Backup Path)” that can be configured via “explicit routing” using “LDP/RSVP signaling.” App. 15 (Owens) at 3:4-12; *see also id.* at 2:40-55; 5:1-7:15; 13:29-45; 9:52-63;

261. Moreover, Pseudowires are set up using the same signaling protocols as regular Pseudowires, regardless of whether they are ultimately designated as primary or standby. The main difference is the particular parameters assigned to the Pseudowire during the set-up process, such as the primary/back-up status or the relative priority.

262. Thus, it would have been obvious to use the method of configuring a Pseudowire disclosed in Hofmeister to configure both regular Pseudowire and standby Pseudowire, particularly because Hofmeister expressly notes the desirability of having a “backup” in at least some circumstances. App. 4 (Hofmeister) at [0137]; [0257]; [0397].

**“determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire”**

263. Claim 1, 9 and 14 recite in part: “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.”

264. To the extent that Hofmeister is deemed to not expressly or inherently disclose this element, it is my opinion that this element would have been obvious in view of RFC 3386 and Owens.

265. As noted above Hofmeister discloses “preempting existing traffic” in the context of the set-up/configuration stage. Under the broadest reasonable interpretation of this claim element as proposed by BSL, wherein “existing traffic” is any “working traffic” and a “priority” is any “preference,” the preemption of traffic flowing on a current Pseudowire when a Pseudowire with a higher “Setup Priority” makes a setup request would disclose this element.

266. RFC 3386 discloses “preempting existing traffic” based on the “relative priority” assigned to the Pseudowire in the context of a network failure. App. 5 (RFC 3386) at § 2.2.2 (“Extra traffic, also referred to as preemptable traffic, is the traffic carried over the protection entity while the working entity is active. Extra traffic is not protected, *i.e.*, ***when the protection entity is required to protect the traffic that is being carried over the working entity, the extra traffic is preempted.***”); § 2.3 (“In the 1:n protection architecture . . . [w]hen multiple working entities have failed simultaneously, only one of them can be restored by the common protection entity. This contention could be resolved by ***assigning a different preemptive priority*** to each working entity.”).

267. Owens also discloses “preempting existing traffic on a standby” based on a “priority” in the context of network failure. For example, Owens teaches that,

in a 1+1 protection scheme, the backup path “could be used to transmit *an exact copy of the working traffic*, with *a selection between the traffic* on the working and protection paths being made at the PML.” App. 15 (Owens) at 6:56-59. Owens further teaches that, in a 1:1 protection scheme, “the working traffic normally travels only on the working path, and is switched to the protection path only when the working entity is unavailable. *Once the protection switch is initiated, all the low priority traffic being carried on the protection path is discarded to free resources for the working traffic.*” *Id.* at 7:1-6; *see also* 5:23-29; 1:34-36 (“[A] protection priority could be used as a differentiating mechanism for premium services.”). Thus, Owens discloses “preempting existing traffic on a standby” under both BSL’s interpretation of this element (*e.g.*, the 1+1 protection scheme) and Petitioner’s interpretation (*e.g.*, the 1:1 protection scheme, with low-priority traffic on the backup during normal operation).

268. Because Hofmeister already discloses assigning relative priorities (Setup/Holding) to a Pseudowire during configuration (which are the same examples of “priority” discussed in the ’652 specification), it would have been an obvious and predictable step for a PHOSITA to use those priorities to make decisions about Pseudowire preemption during a network failure, as taught by RFC 3386 and Owens.

269. Moreover, it would have been highly inefficient and unusual *not* to use the Setup Priority and the Holding Priority to make determinations about preemption during a network failure. Given the costs of running a network, service providers are motivated to make sure that the network is utilized in the most efficient way. As a result, market forces generally cause service providers to ensure that no parts of a network are idle and that all resources and bandwidth are being used as much as possible at any one time. In fact, it is common practice for service providers to over-subscribe a network. Because of this, the use of backup

paths for “extra” or “low priority” traffic was commonplace and “preempting existing traffic” when there is a failure or when the network is oversubscribed is the most reasonable and efficient way to operate the network.

270. In fact, Hofmeister himself supports that a motivation to combine Hofmeister with the concept of “preempting existing traffic” would have existed based on the motivation to make the network more efficient: “By applying pseudo-wire shuffling and preemption techniques, the providers can make a better use of their network resources.” App. 4 (Hofmeister) at Abstract.

271. In addition to RFC 3386, there are *hundreds* of IETF Internet Drafts that discuss the topic of preempting traffic that predated both Hofmeister and the ’652 patent. Some examples can be found in the list attached as Appendix 8. Indeed, “preempting existing traffic” during setup, failure, or oversubscription was a fundamental concept of network communication at the time of Hofmeister. At that time, traffic preemption was used in numerous contexts, such as setting priorities, Quality of Service (QoS), Service Level Agreements (SLA), setting limitations on queues, delays and/or bandwidths, just to name a few.

### **Claims 2-3, 10**

272. Claims 2 and 10 recite in part: “wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.” Claim 10 recites in part: “in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.” Hofmeister does not explicitly disclose these elements. It is my opinion, however, that these element would have been obvious in view of RFC 3386 and Owens.

273. As shown in Section VII(D)(2) and (3), RFC 3386 discloses these elements. App. 5 (RFC 3386) at § 2.2.2, ¶ 7 (“reconfiguration involves switching the affected traffic from a working entity to a protection entity”); *see also* § 2.2.3,

¶ 3 (“When the working entity fails, the protected traffic is switched to the protected entity.”).

274. Owens also discloses these elements. For example, Owens teaches that a “Protection or Backup LSP” is “established to carry the traffic of a Working path (or paths) following failure on the Working path (or one of the Working paths, if more than one) and a subsequent protection switch by the PSL.” App. 15 (Owens) at 3:4-12.

275. As described in connection with claims 1, 9 and 14 above, it would have been obvious to a PHOSITA to use the configuration techniques described in Hofmeister to set up a standby Pseudowire (versus a regular Pseudowire).

276. The purpose of a “standby” is to protect one or more primary or working entities. Moreover, switching traffic from the primary or working entity to the standby or backup upon a failure is a protection technique that was well-known and had been widely used for years. As such, it would have been obvious to use Hofmeister’s technique for configuring Pseudowires to perform the elements described in claims 2-3 and 10.

277. The motivation to combine these references is further evidenced by the fact that Hofmeister repeatedly mentions that it would be desirable to protect the described Pseudowires using, among other things, “a backup link.” Hofmeister further notes that an advantage of the claimed Pseudowire environment is that it can take advantage of the “rich set of features for network resource allocation, traffic *restoration*, and link *protection*” (App. 4 (Hofmeister) at [0257]) and that it allows “*protection* mechanisms” to be “triggered much faster thereby preventing data loss” (*id.* at [0137]). With respect to some of the disclosed embodiments, Hofmeister notes that it would be desirable to direct traffic to a **backup link**. *Id.* at [0397] (“direct existing traffic to a backup link (*e.g.* such as using protection bandwidth triggered via a conventional APS (automatic protection switch) protocol

for SONET/SDH traffic”). Thus, Hofmeister expressly contemplates the use of a primary/standby protection scheme.

278. In fact, the approach taken in Hofmeister is much more sophisticated than a simple primary/backup protection scheme in that it would not only respond to a simple failure, but also employs advanced techniques for bandwidth requirements, failure and restoration by implementing shuffling Pseudowires from one circuit to another based on a comparison of the relevant Setup and Holding Priorities. As such, it is my opinion that Hofmeister assumes that one reading the specification would already know about conventional protection techniques, such as a basic 1:1 or 1+1 protection scheme, and the discussion therein is intended to articulate a more sophisticated protection and bandwidth allocation scheme.

279. In any event, RFC 3386 and Owens would have been an obvious place to look for the specific protection techniques and schemes that could be applied to achieve the protection benefits and efficiency goals that are referenced throughout the Hofmeister specification.

#### **Claim 4**

280. Claim 4 recites: “wherein the standby Pseudowire is dynamically selected from a plurality of connections.” Hofmeister does not explicitly disclose these elements. It is my opinion, however, that this element would have been obvious in view of RFC 3386 and Owens.

281. Based on my review of BSL’s infringement contentions, it is clear that BSL is interpreting “wherein the standby Pseudowire is dynamically selected from a plurality of connections” to encompass situations where a router selects which traffic to accept between a primary and secondary pseudowire that are configured in a 1+1 protection scheme. *See* App. 25 (BSL’s Preliminary Infringement Contentions) at 12 (“The local PE router automatically selects that between the primary and standby pseudowires, traffic from which pseudowire is accepted . . .

which is an embodiment of the standby Pseudowire is dynamically selected from a plurality of connections.”)

282. As shown in Section VII(D)(4), RFC 3386 discloses this element. App. 5 (RFC 3386) at § 2.3 (teaching “dynamic selection and establishment of alternate paths.”); *see also* § 2.2.3. It would have been obvious to a person of skill in the art to combine the Pseudowire signaling techniques disclosed in Hofmeister with RFC 3386 for the same reasons discussed in connection with claims 1, 9 and 14 above. Indeed, the method described in Hofmeister is intended to allow flexibility and agility in setting up Pseudowires, and it would have been obvious to also use that flexibility and agility in connection with responding to network failures.

283. RFC 3386 and Owens also discloses this element under BSL’s apparent claim construction. For example, both disclose a 1+1 protection scheme. *See* claims 5, 8, 11, 13, 15, and 17 below.

**Claims 5, 8, 11, 13, 15 and 17**

284. Claims 5, 11 and 15 recite “wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.” Claims 8, 13, and 17 further recite: “wherein the protection scheme indicates at least one of the following: . . . a 1+1 . . . ; a 1:1 . . . ; a 1:N . . . ; or an M:N protection scheme.”

285. Hofmeister discloses these particular protection schemes. For example, Hofmeister discloses that “protection bandwidth” can be “triggered via conventional APS (automatic protection switch) protocol for SONET/SDH traffic.” App. 4 (Hofmeister) at [0397].

286. It was well-known in the art that “APS protocol” included 1+1, 1:1, 1:N and M:N protection schemes. For example, RFC notes that “[p]rotection techniques can be implemented by several architectures: 1+1, 1:1, 1:n, and m:n. In

the context of SDH/SONET, they are referred to as Automatic Protection Switching (APS).” App. 5 (RFC 3386) at 2.2.3.

287. Owens also discloses these very same protection schemes (App. 15 (Owens) at 6:45 – 7:15).

288. Because the invention disclosed in Hofmeister explicitly draws from other, common concepts that existed in SONET, MPLS and GMPLS (*see* App. 4 (Hofmeister) at [0016]), and because Hofmeister expressly notes that it would be desirable to incorporate the APS protocol from SONET (*id.* at [0397]), it would have been a predictable and natural step for a person skilled in the art to combine these conventional protection schemes with the specific Pseudowire environment disclosed in Hofmeister.

289. In sum, it is my opinion that Hofmeister in view of RFC 3386 and Owens renders each of the Challenged Claims obvious under 35 U.S.C. § 103.



**D. RFC 3386 Anticipates Claims 1-5, 8-11, 13-15, and 17 of the '652 Patent under 35 U.S.C. § 102.**

290. RFC 3386, entitled “Network Hierarchy and Multilayer Survivability” was published by the IETF in November 2002. RFC 3386 is prior art to the '652 patent under § 102(b) as a printed publication, as it was published more than one year before February 14, 2005, the earliest possible priority date for the '652 patent.

291. It is my opinion that RFC 3386 anticipates claims 1-5, 8-11, 13-15 and 17 of the '652 patent under 35 U.S.C. § 102(b).

292. RFC 3386 provides an overview of survivability and hierarchy techniques that can be used in a variety of service provider environments. App. 5 (RFC 3386) at § 1. “Network survivability” is a synonym for “network protection” in this context.

293. While RFC 3386 discusses protection techniques that can be used in a variety of service provider environments, it also specifically identifies *Pseudowires* as one of those environments. For example, the document notes that there are several “pressing needs” with respect to “horizontal hierarchy” in data networks, including the requirement “to set up many Label Switched Paths (LSPs) in a service provider network . . . to support layer 2 and layer 3 Virtual Private Network (VPN) services that require edge-to-edge signaling across a core network.” *Id.* At the time of RFC 3386, this use of LSP tunnels to accomplish “edge-to-edge signaling” in layer 2 and layer 3 VPNs was a well-known technique and was also referred to in some contexts as using a “Pseudowire.” For example, RFC 3386 notes that “[t]here are a number of different approaches to layer 2 and layer 3 VPNs and they are currently being addressed by different emerging protocols in the provider-provisioned VPNs (*e.g.*, virtual routers) and Pseudo Wire

Edge-to-Edge Emulation (PWE3) efforts based on either MPLS and/or IP tunnels.” App. 5 (RFC 3386) at § 4.2.

1. *Claim 1: A method of providing protection to network traffic, comprising,*

294. Claim 1 recites: “A method of providing protection to network traffic.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

295. Indeed, RFC 3386 notes that its primary purpose is to provide requirements for network survivability, which is synonymous with network protection: “This document presents a proposal of the near-term and practical requirements for network survivability and hierarchy in current service provider environments.” App. 5 (RFC 3386) at § 1; *see also* § 2.2.1 (“Survivability is the capability of a network to maintain service continuity in the presence of faults within the network”).

a. *sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire,*

296. Claim 1 further recites: “sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

297. For example, RFC 3386 teaches that the protection techniques that it discloses are intended to apply to a variety of different types of network paths and

entities. For example, in § 2.2.1, RFC 3386 notes that “protection and restoration are implemented either on a per-link basis, on a per-path basis, or throughout an entire network.” RFC 3386 also generally defines “working entity” as “the entity that is used to carry traffic in normal operation mode” and “protection entity” as “the entity that is used to carry protected traffic in recover operation mode, *i.e.*, when the working entity is in error or has failed.” App. 5 (RFC 3386) at § 2.2.2.

298. RFC 3386 specifically identifies Pseudowires as one of the environments in which the disclosed protection techniques can be applied. For example, § 4 teaches that “[e]fforts in the area of network hierarchy should focus on mechanisms that would allow more scalable edge-to-edge signaling, or signaling across networks with existing network hierarchy (such as multi-area OSPF).” One skilled in the art would understand that type of “edge-to-edge” signaling is referring to the concept of Pseudowires.

299. Section 4.1 goes on to explain that “concatenation of survivability approaches can be used to cascade across a horizontal hierarchy.” Further evidence that RFC 3386 applies to Pseudowires can be found in § 4.2. For example, RFC 3386 states: “A primary driver for intra-domain horizontal hierarchy is signaling capabilities in the context of edge-to-edge VPNs . . . currently being addressed by different emerging protocols in the . . . Pseudo Wire Edge-to-Edge Emulation (PWE3) efforts based on either MPLS and/or IP tunnels.” App. 5 (RFC 3386) at § 4.2. RFC 3386 further notes that “most service providers feel that  $O(N^2)$  meshes are not necessary for VPNs, and that the number of tunnels to support VPNs would be within the scalability bounds of current protocols and implementations.” *Id.*

300. In § 5, RFC 3386 further teaches that “approaches as described above that allow concatenation of survivability schemes across hierarchical boundaries seem sufficient.”

301. One skilled in the art would understand that these disclosures in RFC 3386 mean that such protection and restoration techniques are intended to apply to a PWE3 and/or Pseudowire environment, both of which existed at the time of RFC 3386 and which were well-known in the art.

302. Thus, it is my opinion that the more general disclosures in RFC 3386 that refer to a “working entity/path” or “protection entity/path” are understood to include Pseudowires, as well as many other similar types of entities and paths.

303. RFC 3386 further discloses a “standby Pseudowire.” For example, RFC 3386 teaches that a “protection entity” can be established. App. 5 (RFC 3386) at § 3.2.1.

304. RFC 3386 also discloses a “source node” and a “destination node.” For example, RFC 3386 teaches that “the head end of a working connection establishes a protection connection to the destination.” App. 5 (RFC 3386) at § 3.2.1. In this example, the “head end” is the “source node” and the “destination” is the “destination node.”

305. RFC 3386 also teaches sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between the source node and the destination node that indicates a protection property associated with the standby Pseudowire. For example, RFC 3386 teaches that “preemptable traffic may be excluded from local restoration . . . This type of control may require the *definition of an object* in signaling.” *Id.* at 3.2.3. The “object” used during signaling is an example of a “Pseudowire protection configuration parameter.”

306. As another example, RFC 3386 teaches that “[t]here should be the ability to maintain *relative restoration priorities* between working and protection connections, as well as between different classes of protection connections.” *Id.* at § 3.2.1. RFC 3386 further teaches that “the ability to signal that *traffic will be sent on both connections* (1+1 Path for signaling purposes) would be valuable in non-

packet networks,” and that “[s]ome *distinction between working and protection connections* is likely, either through explicit objects, or preferably through implicit methods such as general classes or priorities.” *Id.* at § 3.2.1. To those skilled in the art, “signaling” in this context is used to refer to the process of sending configuration parameters between a source node and destination node to negotiated a connection.

307. It is my understanding that the parties have agreed in the Concurrent Litigation that the term “Pseudowire protection configuration parameter” means “data structure with one or more fields that specify certain protection properties associated with a Pseudowire.” Under that construction, the “definition of an object in signaling,” the “explicit objects” and/or the “implicit methods such as general classes or priorities” that RFC 3386 discloses as being signaled between the head end and the destination comprise a “protection configuration parameter.” Indeed, one skilled in the art would understand that an “object” or “explicit object” is a term used in the art to refer to a data structure that contains fields to hold configuration information, such as the “objects” used in connection with LDP or RSVP-TE protocol.

308. It is my understanding that the parties have agreed in the Concurrent Litigation that the term “protection property” should be construed as “field of data that corresponds to a protection scheme, protection type, domain type, and/or priority.” Under that construction, RFC 3386 discloses that the “protection configuration parameter” includes a “protection property.” For example, the “relative restoration priorities” disclosed by RFC 3386 (§ 3.2.1) are the same or a similar concept to the “priority” concept that is disclosed in the ’652 patent. In addition, the indication “that traffic will be sent on both connections” (*id.*) is the same or a similar concept to the “protection scheme” disclosed in the ’652 patent.

309. As additional examples of protection-related configuration parameters, §§ 2.3 and 3.2 discuss the various protection-related configuration parameters that need to be set to handle network protection) and § 2.2.3 discusses the passing of parameters and information between entities to set and handle fail over and traffic protection of network). In addition, § 3.2.2 discloses the example of setting protection configurations that need to be propagated in the network, which requires pre-signaling or pre-planning. In this context, pre-signaling or pre-planning would require the sending of parameters between nodes in the network.

310. Indeed, the “distinction between working and protection connections” (App. 5 (RFC 3386) at (§ 3.2.1)) is the same or similar to one of the “protection properties” that BSL has accused of infringing the ’652 patent (*i.e.*, a designation as standby vs. primary). As a result, this would be a “protection property” under BSL’s interpretation of the claims.

*b. the protection property including a priority for the standby Pseudowire*

311. Claim 1 further recites: “the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

312. For example, under the construction of “priority” that is advocated by BSL (*i.e.*, “preference”), the “distinction between the working and protection connections” discussed in RFC § 3.2.1 is “a priority” for the standby Pseudowire. Indeed, based on my review of BSL’s infringement contentions, it is clear that BSL is contending that the mere designation of a Pseudowire as “primary” or “backup” is sufficient to satisfy the requirement that the “protection property” include a “priority.”

313. Alternatively, under a narrower construction of “priority,” the “relative restoration priorities” discussed in RFC 3386 § 3.2.1 are an example of “a

priority.” RFC 3386 discloses that the “relative restoration priorities” are maintained between working and protection connections, as well as between different classes of protection connections. *Id.*

314. As another example, RFC 3386 discloses the use of a “restoration priority.” App. 5 (RFC 3386) at § 2.2.4. RFC 3386 teaches that “[r]estoration priority is a method of giving preference to protect higher-priority traffic ahead of lower-priority traffic. Its use is to help determine the order of restoring traffic after a failure has occurred.” *Id.*

315. As another example, RFC 3386 discloses the use of a “preemption priority.” App. 5 (RFC 3386) at § 2.2.4. RFC 3386 teaches that “[p]reemption priority is a method of determining which traffic can be disconnected in the event that not all traffic with a higher restoration priority is restored after the occurrence of a failure.”

316. In order to implement a “restoration” or “preemption” priority, one must necessarily associate a priority with the standby Pseudowire that is being set up, or else there would be no way to perform the actions disclosed – *i.e.*, determine whether a particular path should be restored on one path over another after a failure has occurred, or determine whether traffic on one path can be preempted to allow a higher-priority path to continue sending traffic when there is a failure.

*c. receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

317. Claim 1 further recites: “receiv[ing/e] a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

318. For example, RFC 3386 teaches that the “protection entity” (*i.e.*, the standby Pseudowire) is established between a “source node” and a “destination node”: “the head end of a working connection establishes a protection connection to the destination.” *Id.* at § 3.2.1. Under the broadest reasonable construction of “receiving a Pseudowire configuration acknowledgement” as proposed by BSL—*i.e.*, that “configuration acknowledgement” means “an indication of whether the destination node accepts the standby Pseudowire,” and that “receiving a Pseudowire configuration acknowledgement . . .” includes the mere execution and completion of a configuration command on a source node—the establishment of a protection connection (which inherently requires the execution and completion of a configuration command) would satisfy this element.

319. As another example, RFC 3386 teaches that the “Interior Gateway Protocol (IGP)” can be used to set up many LSPs to support layer 2 and layer 3 VPN services that require edge-to-edge signaling across a core network. App. 5 (RFC 3386) § 1. “Layer 2 and layer 3 VPN services that require edge-to-edge signaling across a core network” is another way of referring to “Pseudowire” technology. “IGP” is a type of protocol used for exchanging routing information between gateways (nodes/routers) within an enterprise network or an Autonomous System. *E.g.*, typically an ISP or a very large organization. Specific examples of IGP protocols include Open Shortest Path First (OSPF), Routing Information Protocol (RIP) and Intermediate System to Intermediate System (IS-IS). Each of these protocols utilizes a “configuration acknowledgement” that indicates whether the requested configuration parameters have been accepted.

320. For example, RFC 2328, entitled “OSPF Version 2” (published in April 1998), discusses “sending link state acknowledgement packets.” App. 16 (RFC 2328) at pg. 152 (“13.5. Sending Link State Acknowledgment packets . . . Each newly received LSA must be acknowledged. This is usually done by sending



Link State Acknowledgment packets. However, acknowledgments can also be accomplished **implicitly** by sending Link State Update packets (see step 7a of Section 13).”).

321. As another example, RFC 1582, entitled “Extensions to RIP to Support Demand Circuits,” (published in February 1994) discloses that “[a]n acknowledgement and retransmission mechanism is provided to ensure that routing updates are received.” App. 17 (RFC 1582) at pg. 1, abstract.

322. As another example, in IS-IS protocol, “[t]he RA bit is sent by the neighbor of a (re)starting router to acknowledge the receipt of a restart TLV with the RR bit set.” App. 18 (RFC 5306) pg. 6 .

323. These RFCs concerning the IGP protocols were standard background knowledge for those of ordinary skill in the art and are being cited as other written evidence solely to help explain RFC 3386.

324. Thus, the disclosure that IGP protocols can be used to signal Pseudowires comprises “receiving an acknowledgement message . . .” under BSL’s apparent claim construction.

*d. accepting the Pseudowire protection configuration parameter by the destination node;*

325. Claim 1 further recites: “accepting the Pseudowire protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

326. For example, RFC 3386 discloses protection techniques where both the working and protection entities are pre-established. App. 5 (RFC 3386) § 2.2.3, ¶ 1 (“as the working entity is established, a protection entity is also established”); § 3.2.1, ¶ 1 (“working connection establishes a protection connection to the destination”), ¶ 2 (capacity in the protection connection is pre-established, however it should be capable of carrying preemptable extra traffic in non-packet

networks); § 3.2.2 , ¶ 1 (“the protection connection in this case is also pre-signaled . . . the mechanism supports the ability for the protection capacity to be shared, or ‘double-booked’”).

327. One skilled in the art would understand that in order for a protection entity to be “established” or “pre-established,” the destination node must accept the configuration parameters.

328. Indeed, under standard routing protocols for setting up virtual paths (such as LDP and RSVP-TE), if a destination node sends a configuration acknowledgement to the source node, it means that the destination node received the message containing the configuration parameters and agreed to it. If the destination node does not accept the parameters, it would send a notification or other response indicating that fact, or it will ignore the Hello message. *See, e.g.*, App. 9 (RFC 3209) at pg. 54, second paragraph (“The Hello extension is fully backwards compatible. The Hello class is assigned a class value of the form 0bbbbbbb. Depending on the implementation, implementations that do not support the extension will either silently discard Hello messages or will respond with an “Unknown Object Class” error. In either case the sender will fail to see an acknowledgment for the issued Hello.”).

329. RFC 3209 was standard background knowledge for those of ordinary skill in the art and is being cited as other written evidence solely to help explain RFC 3386.

330. Thus, any “establishment” of a protection path between the head end and the destination necessary means that the destination node has “accept[ed] the Pseudowire protection configuration parameter” under BSL’s apparent claim construction.

*e. using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

331. Claim 1 further recites: “using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

332. For example, RFC 3386 discloses multiple instances where a protection entity is used after it is set up. *See, e.g.*, App. 5 (RFC 3386) at § 2.2.3 (discussing the operation of protection entities using different protection switching techniques); § 2.2.4 (describing the pros and cons of protection entities that use protection switching as opposed to restoration); § 2.3 (comparing various protection mechanisms); § 3.2.1 (describing operation of 1:1 Path Protection with Pre-Established Capacity); § 3.2.2 (describing operation of 1:1 Path Protection with Pre-Planned Capacity).

333. RFC 3386 further teaches that, where the protection entity has been designated with a 1+1 protection scheme, “[i]n normal operation mode, identical traffic is transmitted simultaneously on both the working and protection entities . . . . A selection between working and protection entity is made based on some predetermined criteria, such as the transmission performance requirements or defect indication.” App. 5 (RFC 3386) at § 2.2.3. Under BSL’s apparent claim construction, this 1+1 operation would comprise “using the standby Pseudowire . . . .”

334. As another example, RFC 3386 discloses that, where the protection entity has been designated with a 1:1 protection scheme, “protected traffic is normally transmitted by the working entity. When the working entity fails, the protected traffic is switched to the protection entity.” *Id.* at § 2.2.3; see also § 3.2.1

(discussing use of protection paths with 1:1 pre-established capacity); § 3.2.2 (discussing use of protection paths with 1:1 pre-planned capacity). Under BSL’s apparent claim construction, this 1:1 operation would comprise “using the standby Pseudowire . . . .”

335. These disclosures constitute “using the standby Pseudowire [*i.e.*, protection entity] that is configured based at least in part on the Pseudowire protection configuration parameter [*i.e.*, that is configured based at least in part on the protection scheme signaled through explicit object or some other implicit means]” under BSL’s apparent claim constructions.

*f. determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

336. Claim 1 further recites: “determining whether to preempt existing traffic on the standby Pseudowire wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

337. For example, RFC 3386 teaches a “1+1 protection architecture” where “[i]n normal operation mode, identical traffic is transmitted simultaneously on both the working and protection entities. At the other end (sink) of the protected domain, both feeds are monitored for alarms and maintenance signals. ***A selection between the working and protection entity is made based on some predetermined criteria***, such as the transmission performance requirements ***or defect indication***.” App. 5 (RFC 3386) at § 2.2.3. Under the broadest reasonable interpretation of this element, as advocated by BSL, the “identical traffic [that] is transmitted simultaneously” would comprise “existing traffic on the standby Pseudowire” and selection between the working and protection entity based on some predetermined

criteria would comprise “preempt[ing] existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire [*i.e.*, the fact that it is designated as a protection entity, not a working entity].” Indeed, under BSL’s infringement theory, any predetermined criteria that indicate which path should be used during normal operation would qualify as a “priority” under BSL’s constructions because it would indicate a “preference” for using that particular path, if it is in working condition.

338. As another example, RFC 3386 teaches a different protection scheme where “[e]xtra traffic, also referred to as preemptable traffic, is the traffic carried over the protection entity while the working entity is active. Extra traffic is not protected, *i.e.*, when the protection entity is required to protect the traffic that is being carried over the working entity, the extra traffic is preempted.” App. 5 (RFC 3386) at § 2.2.2; *see also id.* at § 2.3 (noting that in a 1:1 architecture, “the protection entity can optionally be used to carry low-priority extra traffic in normal operation, if traffic preemption is allowed”). RFC 3386 also teaches that, in some protection schemes, decisions about preemption can be determined based on the priority assigned to the relevant paths: “In the 1:n protection architecture . . . [w]hen multiple working entities have failed simultaneously, only one of them can be restored by the common protection entity. This contention could be resolved by ***assigning a different preemptive priority*** to each working entity. As in the 1:1 case, the protection entity can optionally be used to carry preemptable traffic in normal operation.” *Id.*

339. Under a narrower construction of “preempting existing traffic on the standby Pseudowire” that requires the “existing traffic” to be different from the traffic that is being transmitted on the primary Pseudowire, the “extra traffic” on the “protection entity” would comprise “existing traffic on the standby Pseudowire.” Similarly, under a narrower construction that requires “priority” to

be more than a mere designation as primary vs. backup, the use of a “preemptive priority” to make the decision as to which working entity to protect when multiple entities fail would comprises “preempt[ing] existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.”

340. By way of example, if the principals of RFC 3386 were used to create three Pseudowires: (1) Protection PW A with a “preemptive priority” of 5, (2) Working PW B with a “preemptive priority” of 6 and that is protected by Protection PW A, and (3) Working PW C with a “preemptive priority” of 7 and that is protected by Protection PW A. Under the teaching of RFC 3386, if PWs B and C both fail, the “preemptive priorities” of the PWs would be examined to determine which Pseudowire should be protected. In this instance, PW C would preempt the traffic on PW A because it has the highest “preemptive priority,” thus resulting in “determining whether to preempt existing traffic on the standby Pseudowire . . . based, at least in part, on the priority for the standby Pseudowire.”

341. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 1 of the '652 patent under BSL's apparent claim constructions under BSL's apparent claim constructions, and thus RFC 3386 anticipates claim 1.

2. *Claim 2: A method as recited in claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.*

342. Claim 2 recites: “a method as recited in claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.”

343. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL's apparent claim construction, RFC 3386 discloses the additional elements of claim 2.

344. For example, RFC 33886 discloses a protection scheme where a protection entity (*i.e.*, standby Pseudowire) is configured to provide protection to at least one working entity (*i.e.*, primary Pseudowire). *See, e.g.*, App. 5 (RFC 3386) at § 2.2.3 (discussing 1+1, 1:1 and 1:n protection schemes).

345. More specifically, in the 1+1 protection scheme disclosed by RFC 3386, a standby Pseudowire is set up to protect a primary Pseudowire and the protected traffic is sent over both the standby and the primary. App. 5 (RFC 3386) at § 2.2.3. If the primary Pseudowire fails, then the system begins accepting traffic from the standby Pseudowire instead of the primary Pseudowire. *Id.* In the 1:1 protection scheme disclosed by RFC 3386, a standby Pseudowire is set up to protect a primary Pseudowire, but the protected traffic is only sent over the primary Pseudowire during normal operation. *Id.* If the primary Pseudowire fails, then the protected traffic is switched over to the standby Pseudowire. *Id.* In a 1:n protection scheme, a standby Pseudowire is set up to protect multiple primary Pseudowires. *Id.* The protected traffic is only sent over the primary Pseudowires during normal operation. *Id.* If one or more of the primary Pseudowires fail, the protected traffic is switched over to the standby Pseudowire. *Id.* Under BSL's apparent constructions, each of these scenarios would disclose claim 2. *Id.*

346. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 2 of the '652 patent under BSL's apparent claim constructions, and thus RFC 3386 anticipates claim 2.

3. *Claim 3: A method as recited in claim 1 wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.*

347. Claim 3 recites: “A method as recited in claim 1 wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.”

348. I incorporate by reference the portions of this declaration pertaining to Claims 1 and 2 above. Under BSL’s apparent claim construction, RFC 3386 discloses the additional elements of claim 2.

349. For example, RFC 3386 teaches that, when the working entity fails to transfer network traffic, the data is switched from the working entity to the protection entity. *See, e.g.*, App. 5 (RFC 3386) at § 2.2.2, ¶ 7 (“In protection, reconfiguration involves switching the affected traffic from a working entity to a protection entity.”); *see also* § 2.2.3, ¶ 3 (“When the working entity fails, the protected traffic is switched to the protected entity.”).

350. More specifically, in the 1+1 protection scheme disclosed by RFC 3386, a standby Pseudowire is set up to protect a primary Pseudowire and the protected traffic is sent over both the standby and the primary. App. 5 (RFC 3386) at § 2.2.3. If the primary Pseudowire fails, then the system begins accepting traffic from the standby Pseudowire instead of the primary Pseudowire. *Id.* In the 1:1 protection scheme disclosed by RFC 3386, a standby Pseudowire is set up to protect a primary Pseudowire, but the protected traffic is only sent over the



primary Pseudowire during normal operation. *Id.* If the primary Pseudowire fails, then the protected traffic is switched over to the standby Pseudowire. *Id.* In a 1:n protection scheme, a standby Pseudowire is set up to protect multiple primary Pseudowires. *Id.* The protected traffic is only sent over the primary Pseudowires during normal operation. *Id.* If one or more of the primary Pseudowires fail, the protected traffic is switched over to the standby Pseudowire. *Id.* Under BSL's apparent constructions, each of these scenarios would disclose claim 3. *Id.*

351. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 3 of the '652 patent under BSL's apparent claim constructions, and thus RFC 3386 anticipates claim 3.

4. *Claim 4: A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.*

352. Claim 4 recites: "A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections."

353. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL's apparent claim construction, RFC 3386 discloses the additional elements of claim 4.

354. For example, RFC 3386 teaches that "dynamic selection and establishment of alternate paths" is possible. *See, e.g.,* § 2.3 (discussing restoration techniques in which "the time it takes for the dynamic selection and establishment of alternate paths may vary").

355. In addition, RFC 3386 teaches a 1+1 protection scheme. App. 5 (RFC 3386) at § 2.2.3. Under BSL's apparent construction of this claim, which encompasses situations where "[t]he local PE router automatically selects that between the primary and standby pseudowires, traffic from which pseudowire is

accepted,” (*see* App. 25 (’652 BSL Preliminary Infringement Contentions) at 12), these schemes disclose claim 4.

356. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 4 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 4.

5. *Claim 5: A method as recited in claim 1, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

357. Claim 5 recites: “A method as recited in claim 1, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.”

358. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL’s apparent claim construction, RFC 3386 discloses the additional elements of claim 5.

359. For example, RFC 3386 discloses that the “protection configuration parameter” can include a “protection scheme” under BSL’s apparent claim construction. For example, RFC 3386 teaches that there should be an “ability to signal that traffic will be sent on both connections (1+1 Path for signaling purposes).” App. 5 (RFC 3386) at § 3.2.1. It would be apparent to one skilled in the art that “signaling” this information between the source and destination would require the information to be included in some sort of configuration parameter.

360. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 5 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 5.

8. *Claim 8: A method as recited in claim 5, wherein the protection scheme indicates at least one of the following:*

361. Claim 8 recites: “A method as recited in claim 5, wherein the protection scheme indicates at least one of the following.”

362. I incorporate by reference the portions of this declaration pertaining to Claims 1 and 5 above. Under BSL’s apparent claim construction, RFC 3386 discloses the additional elements of claim 8.

*a. a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

363. Claim 8 further recites: “a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

364. For example, RFC 3386 discloses a 1+1 protection scheme. *See, e.g.*, App. 5 (RFC 3386) at § 2.2.3, ¶ 2.

*b. a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

365. Claim 8 further recites: “a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

366. For example, RFC 3386 discloses a 1:1 protection scheme. *See, e.g.*, App. 5 (RFC 3386) § 2.2.3, ¶ 3.

*c. a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or*

367. Claim 8 further recites: “a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

368. For example, RFC 3386 discloses a 1:N protection scheme. *See, e.g.*, App. 5 (RFC 3386) at § 2.2.3, ¶ 4.

*d. an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.*

369. Claim 8 further recites: “an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

370. For example, App. 5 (RFC 3386) at § 2.2.3, ¶ 5. *See, e.g.*, App. 5 (RFC 3386) at § 2.2.3, ¶ 3

371. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 8 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 8.

**Claims 9-11, 13-15 and 17**

372. I note that the limitations of independent claim 9 and claim 14 are nearly identical to claim 1, except for the fact that claim 9 is a system claim and claim 14 is a computer program claim. The dependent claims also mirror the claims that depend on claim 1. For example, claim 10 is analogous to claim 2, claims 11 and 15 are analogous to claim 5, and claims 13 and 17 are analogous to claim 8. As such, my analysis below largely incorporates by reference my analysis with respect to claims 1-5 and 8. The claims correlate as follows:

<b><u>Method</u></b>	<b><u>System</u></b>	<b><u>Computer Program</u></b>
Claim 1	Claim 9	Claim 14
Claim 2	Claim 10	
Claim 3		
Claim 4		

Claim 5	Claim 11	Claim 15
Claim 8	Claim 13	Claim 17

9. *Claim 9: A system for providing protection to network traffic, comprising: a processor configured to:*

373. Claim 9 recites: “a system for providing protection to network traffic, comprising a processor.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

374. I incorporate by reference my comments from claim 1 above.

375. In addition, it would be apparent to one skilled in the art that a processor would be required in order to perform the protection techniques described in RFC 3386.

*a. send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;*

376. Claim 9 recites: “send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

377. I incorporate by reference my comments from claim element 1(a) and 1(b) above.

- b. receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

378. Claim 9 recites: “receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

379. I incorporate by reference my comments from claim element 1(c) above.

- c. accept the Pseudowire protection configuration parameter by the destination node;*

380. Claim 9 recites: “accept the Pseudowire protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

381. I incorporate by reference my comments from claim element 1(d) above.

- d. use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

382. Claim 9 recites: “use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

383. I incorporate by reference my comments from claim element 1(e) above.

- e. determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based,*

*at least in part, on the priority for the standby Pseudowire.*

384. Claim 9 recites: “determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

385. I incorporate by reference my comments from claim element 1(f) above.

386. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 9 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 9.

*10. Claim 10: A system as recited in claim 9, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.*

387. Claim 10 recites: “A system as recited in claim 9, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this claim.

388. I incorporate by reference my comments from claims 1, 2, and 9 above.

389. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 10 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 10.

11. *Claim 11: A system as recited in claim 9, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

390. Claim 11 recites: “wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.” Under BSL’s apparent claim construction, RFC 3386 discloses this claim.

391. I incorporate by reference my comments from claims 1, 5 and 9 above.

392. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 11 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 11.

13. *Claim 13: A system as recited in claim 11, wherein the protection scheme indicates at least one of the following:*

393. Claim 13 recites: “A system as recited in claim 11, wherein the protection scheme indicates at least one of the following.” Under BSL’s apparent claim construction, RFC 3386 discloses this claim.

394. I incorporate by reference my comments from claims 1, 5, 9 and 11 above.

*a. a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

395. Claim 13 further recites: “a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

396. I incorporate by reference my comments from claim 8 above.



*b. a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

397. Claim 13 further recites: “a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

398. I incorporate by reference my comments from claim 8 above.

*c. a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or*

399. Claim 13 further recites: “a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

400. I incorporate by reference my comments from claim 8 above.

*d. an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.*

401. Claim 13 further recites: “an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

402. I incorporate by reference my comments from claim 8 above.

403. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 13 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 13.

14. *Claim 14: A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for:*

404. Claim 14 recites: “A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

405. I incorporate by reference my comments from claims 1 and 9 above.

406. In addition, it would be apparent to one skilled in the art that the protection techniques described in RFC 3386 would be implemented using a computer program product that is embodied in a computer readable storage medium and that comprises computer instructions.

*a. sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;*

407. Claim 14 further recites: “sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the

protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

408. I incorporate by reference my comments from claim elements 1(a) and 1(b) above.

*b. receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

409. Claim 14 further recites: “receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

410. I incorporate by reference my comments from claim element 1(c) above.

*c. accept the Pseudowire protection configuration parameter by the destination node;*

411. Claim 14 further recites: “accept the Pseudowire protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

412. I incorporate by reference my comments from claim element 1(d) above.

*d. using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

413. Claim 14 further recites: “using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration

parameter.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

414. I incorporate by reference my comments from claim element 1(e) above.

*e. determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

415. Claim 14 further recites: “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

416. I incorporate by reference my comments from claim element 1(f) above.

417. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 14 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 14.

*15. Claim 15: A computer program product as recited in claim 14, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

418. Claim 15 recites: “A computer program product as recited in claim 14, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

419. I incorporate by reference my comments from claim 1, 5 and 14 above.

420. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 15 of the '652 patent under BSL's apparent claim constructions, and thus RFC 3386 anticipates claim 15.

17. *Claim 17: A computer product as recited in claim 15, wherein the protection scheme indicates at least one of the following:*

421. Claim 17 recites: "A computer product as recited in claim 15, wherein the protection scheme indicates at least one of the following." Under BSL's apparent claim construction, RFC 3386 discloses this element.

422. I incorporate by reference my comments from claims 1, 5, 8, 14, and 15 above.

a. *a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

423. Claim 17 further recites: "a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires." Under BSL's apparent claim construction, RFC 3386 discloses this element.

424. I incorporate by reference my comments from claim 8 above.

b. *a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

425. Claim 17 further recites: "a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire." Under BSL's apparent claim construction, RFC 3386 discloses this element.

426. I incorporate by reference my comments from claim 8 above.

c. *a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or*

427. Claim 17 further recites: "a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires." Under BSL's apparent claim construction, RFC 3386 discloses this element.

428. I incorporate by reference my comments from claim 8 above.

*d. an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.*

429. Claim 17 further recites: “an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

430. I incorporate by reference my comments from claim 8 above.

431. In sum, it is my opinion that RFC 3386 discloses each and every element of claim 17 of the ’652 patent under BSL’s apparent claim constructions, and thus RFC 3386 anticipates claim 17.

432. Thus, it is further my opinion that RFC 3386 anticipates each of the Challenged Claims of the ’652 patent.

**E. RFC 3386 in view of RFC 3209 renders Claims 1-5, 8-11, 13-15, and 17 obvious under 35 U.S.C. § 103**

433. If certain aspects recited in claims 1-5, 8-11, 13-15, and 17 are not deemed to be disclosed or inherent over RFC 3386 alone, then it is my opinion that the inclusion of those aspects certainly would be obvious over RFC 3386 in view of RFC 3209.

434. As described above, RFC 3386 is a document published by the IETF that describes various protection configuration parameters that can be used in a wide range of networks, including SONET, MPLS, GMPLS and Pseudowire environments.

435. RFC 3209 was also published by the IETF no later than December 2001 and is entitled “RSVP-TE: Extensions to RSVP for LSP Tunnels.” RSVP-TE is a protocol that can be used to set up LSP Tunnels, Traffic Engineered Tunnels or Pseudowires. RFC 3209 describes extensions to the RSVP protocol. It allows additional parameters to be associated with the path during configuration so that a provider can have greater control over the traffic flow, bandwidth, and failure.

436. RFC 3209 and RFC 3386 are both in the very narrow field of L2-L3 internet protocols. They also derive from the same tight-knit IETF standards community that was working together and collaborating on issues pertaining to MPLS, traffic engineering, and Pseudowire technologies.

437. Moreover, RFC 3209 and RFC 3386 are both intended to address the same problem—*i.e.*, obtaining greater control over configuration, management and resiliency of network environments that employ virtual paths.

438. The close connection between RFC 3209 and RFC 3386 can be seen from the references themselves. For example, RFC 3386 notes that “[a] primary driver for intra-domain horizontal hierarchy is **signaling** capabilities in the context of edge-to-edge VPNs, potentially across **traffic-engineered** data networks.”

App. 5 (RFC 3386) at § 4.2. Given that RFC 3209 was a widely-known and accepted protocol for signaling and traffic engineering, it would have been an obvious place to look for the specific details regarding the signaling protocols and traffic engineering that could be used to achieve the desired signaling and traffic engineering goals discussed in RFC 3386.

**Claims 1, 9, 14**

439. If certain aspects recited in claims 1, 9 and 14 are not deemed to be disclosed or inherent over RFC 3386 alone, it is my opinion that the inclusion of those aspects certainly would be obvious over RFC 3386 in view of RFC 3209.

**“send[ing] a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property for the standby Pseudowire, the protection property including a priority for the standby Pseudowire”**

440. Claims 1, 9 and 14 recite in part: “send[ing] a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property for the standby Pseudowire, the protection property including a priority for the standby Pseudowire.”

441. To the extent that RFC 3386 is deemed to not expressly or inherently disclose this element, it is my opinion that this element would have been obvious in view of RFC 3209.

442. For example, RFC 3209 teaches that “[t]o create an LSP tunnel, the first MPLS node on the path – that is, the *sender node* with respect to the path – creates an *RSVP Path message* . . . and inserts a LABEL\_REQUEST object into the Path message.” App. 9 (RFC 3209) at § 2.2. RFC 3209 further teaches that “a SESSION\_ATTRIBUTE object can be added to Path messages to aid in session



identification and diagnostics” and “[a]dditional control information, *such as setup and hold priorities, resource affinities . . . and local-protection, are also included in this object.*” *Id.* Thus, RFC 3209 teaches the sending of a protection configuration parameter (*e.g.*, RSVP Path message with “SESSION\_ATTRIBUTE”) that indicates various protection properties (*e.g.*, setup and hold priorities, resource affinities and local protection) between a source node and a destination node. More particularly, RFC 3209 teaches that the “protection property” can include a “priority” (*e.g.*, setup and hold priorities).

443. This signaling process described by RFC 3209 was the standard process for setting up LSP Tunnels, Pseudowires, and other virtual tunnels at the time of RFC 3386 and RFC 3209.

444. In several instances, RFC 3386 notes that “signaling” is used to establishing the new paths (*see, e.g.*, § 2.2.3) and that “[t]here should be the ability to maintain relative restoration priorities between the working and protection connections, as well as between different classes of protection connections.” Thus, it would have been apparent to one skilled in the art that the particular protection configuration parameters discussed in RFC 3386 could be combined with the well-known techniques for signaling new Pseudowires, (including the use of the “Path message” and “SESSION\_ATTRIBUTE” to signal protection configuration parameters) that are disclosed in RFC 3209. Indeed, this would have been nothing more than applying a known technique (RSVP-TE signaling) to a known network environment (Pseudowires) to achieve the predictable result of communicating protection configuration parameters for the purpose of setting up a virtual path in a network.

**receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;**

445. Claims 1, 9 and 14 recite in part: “receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.”

446. To the extent that RFC 3386 is deemed to not expressly or inherently disclose this element, it is my opinion that this element would have been obvious in view of RFC 3209.

447. Under BSL’s apparent claim construction, RFC 3209 discloses this element in great detail: “The destination node of a label-switched path responds to a LABEL\_REQUEST by including a LABEL object in its response RSVP Resv message.” App. 9 (RFC 3209) at § 2.2; *see also* § 4.2.5 (noting that destination noted sends a “PathErr” message with an error code if it cannot accept the requested parameters).

448. Moreover, the use of a configuration acknowledgement message was employed by most (if not all) standard signaling protocols that were used to establish LSP tunnels and Pseudowires, such as LDP, draft-Martini, and even the Hofmeister method described above. For additional details regarding these other protocols, I incorporate by reference my comments above in Section VII(B)(a)(c) and (C) regarding this element as it relates to Hofmeister.

449. Thus, adding a configuration acknowledgement to RFC 3386 would have been nothing more than applying the known technique of using an acknowledgement message to the known protection methods described in RFC 3386 to yield the predictable result of reliable and standardized signaling.

450. Moreover, the Patent Owner did not contest the obviousness of combining an acknowledgement message with methods for configuring a standby

Pseudowire or other standby path when traversing the Examiner's rejections. *See, e.g.,* App. 3 ('652 File History) at pgs. 151-52, 170-74. This is not surprising because, as discussed above, the use of an acknowledgement is a fundamental part of nearly any signaling protocol that was in use at the time of RFC 3386, RFC 3209 and the '652 patent.

451. In sum, it is my opinion that RFC 3386 in view of RFC 3209 renders each of the Challenged Claims obvious under 35 U.S.C. § 103.

**F. Halabi anticipates Claims 1-5, 8-11, 13-15, and 17 under 35 U.S.C. § 102.**

452. As described above, Halabi is a book called “Metro Ethernet” that was published in 2003.

453. It is my opinion that Halabi anticipates claims 1-5, 8-11, 13-15, and 17 of the ’652 patent under 35 U.S.C. § 102(b).

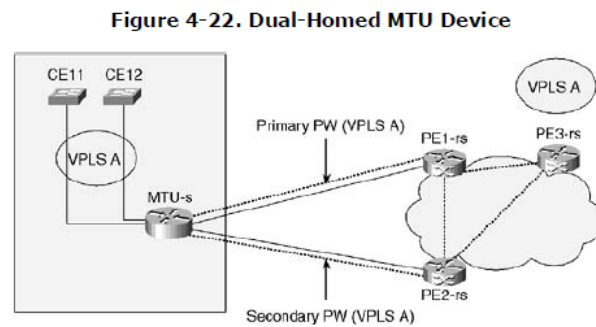
454. Halabi is a part of the Cisco Press series of books, which includes a variety of industry text books and treatises that summarize the state of the art on particular topics. The Cisco Press books generally discuss mainstream topics that are part of the curriculum for networking classes and used also as study guides for Cisco Certifications. The Cisco Press books are also used by engineers who use Cisco products to learn about the available features and configuration options.

455. Halabi in particular provides a detailed discussion of techniques for deploying Ethernet in the Metro, which “requires the scalability and robustness features that exist only in IP and Multiprotocol Label Switching (MPLS) control planes.” App. 6 (Halabi) at pg. xv. Halabi notes that, to deploy Ethernet in the Metro, “hybrid Layer 2 (L2) and Layer 3 (L3) IP and MPLS networks have emerged as a solution that marries Ethernet’s simplicity and cost effectiveness with the scale of IP and MPLS networks.” *Id.* Halabi first discusses how to create these hybrid networks, which employ Pseudowires (*see* Chapter 4), and then discusses how various traffic engineering, fast reroute, and GMPLS protocols can be used to increase the control, reliability and protection of such hybrid networks (*see* Chapters 5-8).

1. *Claim 1: A method of providing protection to network traffic, comprising:*

456. Claim 1 recites: “a method of providing protection to network traffic.” Under BSL’s apparent claim construction, RFC 3386 discloses this element.

457. For example, Halabi depicts a method for providing protection to a Primary Pseudowire in Figure 4-22:



458. In this figure, the “Secondary PW” is set up to provide protection to the “Primary PW.” Additional, detailed information regarding this “dual-homing” protection technique for Pseudowires can be found on pages 102-103 of Halabi.

459. As another example, Halabi discloses the use of RSVP-TE and GMPLS protocols to provide protection to network traffic in case of a failure of a Pseudowire using multiple fault detection and restoration mechanisms. *See id.* at pg. 134, ¶ 1 (“In this chapter, you see how MPLS, through the use of RSVP-TE, **can be used to establish backup paths in the case of failure.**”); Ch. 6, generally.

460. As another example, pages 188-189 of Halabi discuss various protection and restoration mechanisms for Pseudowires and other types of virtual tunnels that can be implemented using the GMPLS protocol. App. 6 (Halabi) at pgs. 188-189. For example, GMPLS provides for “fault detection” using mechanisms such as loss of light, optical signal-to-noise ratio, bit error rate, and Alarm Indicator Signal. *Id.* at pg. 188. This portion of Halabi also teaches that GMPLS offers “fault isolation” using and LMP fault management procedure and “fault notification” using “RSVP-TE Notify messages.” *Id.* Halabi also notes that GMPLS uses the following protection mechanisms: (1) 1+1 protection, (2) 1:1 protection, (3) M:N protection, (4) span protection, (5) span restoration, (6) path protection, and (7) path restoration. *Id.* at pg. 189.

- a. *sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire;*

461. Claim 1 further recites: “sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

462. For example, Halabi teaches Pseudowires and Pseudowire protection. Indeed, Halabi includes an extensive discussion of various methods for emulating Ethernet over a packet-switched network by creating Pseudowires. *See* App. 6 (Halabi) at Chapter 4 (discussing hybrid L2 and L3 IP/MPLS Networks that allow Ethernet to be emulated over MPLS networks using Label-Switched Path (LSP) tunnels).

463. As another example, Halabi teaches that Pseudowires can be set up within LSP tunnels and that protection of the Pseudowires, LSP tunnels and/or traffic trunks can be accomplished using traffic engineering techniques. *See, e.g., id.* at Ch. 5 and Ch. 8. More specifically, Halabi teaches that “When traffic moves from one site to another across the carrier’s backbone [in MPLS L2VPN], it follows the MPLS label switched path (LSP) assigned for that traffic. The LSP itself could have been formed via dynamic routing calculated by the routing protocols. On the other hand, **the LSP could be traffic-engineered** to allow certain types of traffic to follow a well-defined trajectory. Also, **many**

**mechanisms can be used for traffic rerouting in case of failure.** The mechanism used depends on whether the carrier requires normal IP routing or MPLS fast reroute mechanism.” *Id.* at 80. Halabi further discloses that these protection techniques can be implemented using the RSVP-TE and/or GMPLS protocols. *Id.* at pg. 117.

464. In addition, Halabi discloses that “Pseudowire protection configuration parameters” are sent between a source node and a destination node in order to configure the Pseudowire.

465. For example, Halabi discloses that “[t]o establish an LSP tunnel, the ingress LSR sends a PATH message to the egress LSR.” *Id.* at pg. 137, ¶ 3; *see also* pg. 135-139 (discussing generally how LSP tunnels are set up and how parameters are exchanged between nodes on a network using RSVP-TE protocol).

466. In addition, Halabi teaches that the PATH Message can include “several different RSVP objects” (*id.* at pg. 141, ¶ 3), one of which is the “SESSION\_ATTRIBUTE object” that “allows RSVP-TE to set different LSP priorities, preemption, and fast-reroute features” that are used in case of a failure. App. 6 (Halabi) at pg. 144, ¶ 5; *see also id.* at pgs. 140-145 (discussing details of PATH Message and various objects that can be included therein).

467. It is my understanding that the parties have agreed in the Concurrent Litigation that the term “Pseudowire protection configuration parameter” means “data structure with one or more fields that specify certain protection properties associated with a Pseudowire.” Under that construction, the “Path Message” that contains the “SESSION\_ATTRIBUTE” comprises a “Pseudowire protection configuration parameters.” Indeed, the “SESSION\_ATTRIBUTE” disclosed in Halabi is a data structure that contains multiple fields that specify certain protection properties, such as priorities, preemption and fast-reroute features, that are associated with the Pseudowire that is being set up using the

SESSION\_ATTRIBUTE. And, according to Halabi, the SESSION\_ATTRIBUTE message is sent between an ingress LSR (“source node”) and an egress LSR (“destination node”).

468. It is also my understanding that the parties have agreed in the Concurrent Litigation that the term “protection property” should be construed as “field of data that corresponds to a protection scheme, protection type, domain type, and/or priority.” Under this construction, the “**priorities**” “**preemption**” and “**fast-reroute**” features are examples of a “**protection property**” that is included in the Pseudowire protection configuration parameter (*i.e.*, the “Path Message” and/or “SESSION\_ATTRIBUTE”). Indeed the “priorities” are an example of a “priority” protection property; and the “preemption” and “fast-reroute” are examples of a “protection type” or “protection scheme” protection property.

469. Thus, Halabi discloses sending a Pseudowire protection configuration parameter (PATH Message) between a source node (ingress LSR) and a destination node (egress LSR), the Pseudowire protection configuration parameter indicating a protection property (*e.g.*, LSP priorities, preemption, and fast-reroute features).

470. As another example, Halabi teaches additional “protection configuration parameters” that can be used in connection with configuring Pseudowires. For example, Halabi teaches that MPLS traffic engineering can utilize a “**priority attribute**” that “defines the relative importance of traffic trunks” and “determine[s] which paths should be used versus other paths at connection establishment and under fault scenarios” and a “**preemption attribute**” that “determines whether a traffic trunk can preempt another traffic trunk from a given path” and “can be used to ensure that high-priority traffic can always be routed in favor of lower-priority traffic that can be preempted.” *Id.* at pg. 128 ¶¶ 4, 5. These “attributes” are additional examples of “configuration options” that



are used to configure a standby Pseudowire. Moreover, one skilled in the art would understand that “attributes” is a term of art used to refer to parameters or “objects” that comprise data structures that store information about the particular parameters for the relevant Pseudowire that are sent between nodes at the time of set-up in order to configure a Pseudowire according to the appropriate parameters.

471. As another example, Halabi teaches that various enhancements to the traditional MPLS signaling can be made using GMPLS, including various enhancements to “protection information.” *Id.* at pgs. 177, 184. More specifically, Halabi teaches that “GMPLS uses a new object type length value (TLV) field to carry LSP protection information . . . Protection information indicates the LSP’s link protection type . . . Protection information also indicates whether the LSP is a primary or secondary LSP. A secondary LSP is a backup to a primary LSP.” *Id.* at pg. 184, ¶ 5-6. A “TLV field” is a commonly-used term in the art that refers to a data structure that carries configuration parameters for Pseudowires. Because Halabi discloses that various protection-related information can be added to the TLV when using the GMPLS protocol, this is an alternate example of “Pseudowire protection configuration parameters” that include a “protection property” under any possible construction of “Pseudowire protection configuration parameters” and “protection property” that are disclosed in Halabi.

*b. the protection property including a priority for the standby Pseudowire;*

472. Claim 1 further recites: “the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

473. For example, Halabi teaches that the “SESSION\_ATTRIBUTE” that is carried in the PATH message includes fields such as Setup and Holding Priority,

which affect whether a session can preempt or can be preempted by other sessions. App. 6 (Halabi) at pg. 144, ¶ 5.

474. It is my understanding that the Patent Owner has asserted that “priority” be interpreted to mean “preference” and has interpreted “priority” to include the mere designation of a Pseudowire as “primary” or “secondary.” Under BSL’s broad construction, the “Setup Priority” and the “Holding Priority” disclosed by Halabi are more than sufficient to disclose the “priority” element of the ’652 claims. Indeed, “setup priority” and “holding priority” are expressly provided as examples of “a priority” in the specification of the ’652 patent. *See, e.g.*, App. 2 (’652 patent) at Fig. 5; 7:6-25; 6:8-12 (“Pseudowire protection configuration parameter 400 includes four fields: protection scheme, protection type, domain type, and priority. A field may have one or more subfields. For example, the priority field is shown to include a holding priority and a setup priority.”)

475. As another example, Halabi discloses a “FLOW\_SPEC Object” that specifies a desired QoS” using a “numeric parameter” called an “Rspec.” App. 6 (Halabi) at pg. 145. Halabi further teaches that the “FLOW\_SPEC” is used with the SESSION object to define a “flow” to “receive the QoS defined by the flowspec.” *Id.* And, sessions that “do not match any of the filter specs” are preempted as “best-effort traffic.” *Id.*

476. As another example, Halabi also discloses a “priority attribute” that “defines the relative importance of traffic trunks” and “determine[s] which paths should be used versus other paths at connection establishment and *under fault scenarios.*” App. 6 (Halabi) at pg. 128, ¶ 4. The “priority attribute” is another example of “a priority” that is disclosed by Halabi under BSL’s apparent construction.

477. As another example, Halabi also discloses a “preemption attribute” that “determines whether a traffic trunk can preempt another traffic trunk from a given path. Preemption can be used to ensure that high-priority traffic can always be routed in favor of lower-priority traffic that can be preempted. Service providers can use this attribute to offer varying levels of service. A service that has preemption could be priced at a higher rate than a regular service.” *Id.* at pg. 128. It would be apparent to one skilled in the art that a “preemption” parameter is another form of assigning a “priority” to the relevant traffic flow.

478. As another example, Halabi also discloses a “resilience attribute” that “determines the behavior of a traffic trunk when fault conditions occur along the path through which the traffic trunk traverses. The resiliency attribute indicates whether to reroute or leave the traffic trunk as is under a failure condition. More extended resilience attributes could specify detailed actions to be taken under failure, such as the use of alternate paths, and specify the rules that govern the selection of these paths.” *Id.* at pg. 128. It would be apparent to one skilled in the art that a “resilience” parameter is another form of assigning a “priority” to the relevant traffic flow.

479. As another example, Halabi also discloses a “policing attribute” that “determines the actions that should be taken by the underlying protocols when a traffic trunk exceeds its contract as specified in the traffic parameters. Policing is usually done on the input of the network, and it indicates whether traffic that does not conform to a certain SLA should be passed, rate limited, dropped, or marked for further action.” *Id.* at pg. 129. It would be apparent to one skilled in the art that a “policing” parameter based on the terms of an SLA is another form of assigning a “priority” to the relevant traffic flow.

480. As another example, Halabi teaches that “GMPLS uses a new object type length value (TLV) field to carry LSP protection information . . . Protection

information indicates the LSP's link protection type . . . Protection information also indicates whether the LSP is a primary or secondary LSP. A secondary LSP is a backup to a primary LSP.” App. 6 (Halabi) at pg. 184, ¶ 5-6. Under the Patent Owner's broad construction of “priority,” which includes the mere designation of a Pseudowire as primary or secondary/backup, the “protection information [that] indicates whether the LSP is a primary or secondary LSP” is yet another example of a “priority” under BSL's proposed construction.

*c. receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

481. Claim 1 further recites: “receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.” Under BSL's apparent claim construction, Halabi discloses this element.

482. For example, Halabi teaches: “To establish an LSP tunnel, the ingress LSR sends a PATH message to the egress LSR, **which in turn replies with a reservation message (RESV)**. Upon completion of the handshake, an LSP tunnel is established.” App. 6 (Halabi) at pg. 137, ¶ 3.

483. Further to this example, Halabi teaches that “[a]n RESV message is transmitted from the egress LSR toward the ingress in response to the receipt of a PATH message” and that it is used for “distributing label bindings, requesting resource reservations along the path, and specifying the reservation style.” App. 6 (Halabi) at pg. 145 ¶ 5; *see also id.* at pgs. 137-141 (discussing generally how LSP tunnels are set up and parameters are exchanged between nodes on a network using RSVP-TE protocol) and 145-146 (“Details of the RESV Message”). The RESV message is thus a “Pseudowire configuration acknowledgement” that indicates

whether the parameters in the PATH message and SESSION\_ATTRIBUTE (Pseudowire protection configuration parameter) have been accepted by the egress LSR (destination node) under BSL's apparent construction.

*d. accepting the Pseudowire protection configuration parameter by the destination node;*

484. Claim 1 further recites: "accepting the Pseudowire protection configuration parameter by the destination node." Under BSL's apparent claim construction, Halabi discloses this element.

485. For example, Halabi teaches: "To establish an LSP tunnel, the ingress LSR sends a PATH message to the egress LSR, which in turn replies with a reservation message (RESV). **Upon completion of the handshake, an LSP tunnel is established.**" App. 6 (Halabi) at pg. 137, ¶ 3. Thus, the "completion of the handshake" means that the egress LSR has accepted the Pseudowire configuration parameters, thus allowing the path to be established. *See also id.* at pgs. 137-141 (discussing generally how LSP tunnels are set up and parameters are exchanged between nodes on a network using RSVP-TE protocol) and 177-179 (discussing generally enhancements to MPLS signaling protocols when using GMPLS).

486. As another example, when discussing the "Protection Information" that is exchanged in the GMPLS protocol, Halabi teaches that "the connection request is processed only if the desired protection type can be honored." App. 6 (Halabi) at pg. 184 ¶ 4. Thus, the processing of the request necessarily requires that the egress node accept the Pseudowire protection configuration parameters under BSL's apparent construction.

487. Accordingly, Halabi teaches, at least in some instances, "accepting the Pseudowire protection configuration parameter by the destination node."

- e. *using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

488. Claim 1 further recites: “using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter.” Under BSL’s apparent claim construction, Halabi discloses this element.

489. For example, Halabi teaches using the Pseudowire based on the protection properties included in the SESSION\_ATTRIBUTE object “to select alternate LSPs in case of failure in the network.” App. 6 (Halabi) at pg. 144 ¶ 5. Halabi further teaches that specific Setup Priority and Holding Priority that have been assigned to a particular Pseudowire during configuration via the SESSION\_ATTRIBUTE object can affect whether the session can preempt or can be preempted by other sessions during operation. *Id.*

490. As another example, Halabi teaches that in one implementation, “[t]he resources allocated for a secondary LSP may be used by other LSPs until the primary LSP fails over to the secondary LSP. At that point, any set of LSPs that are using the resources for the secondary LSP must be preempted.” App. 6 (Halabi) at pg. 184 ¶ 6.

491. Halabi also teaches the use of a Pseudowire in accordance with various protection properties contained in the protection-related TLV that is disclosed in connection with GMPLS protocol. For example, Halabi discloses how a Pseudowire can be used in accordance with various protection schemes, including 1+1 protection, 1:1 protection, M:N protection schemes, which use (at least in part) “information [that] indicates the LSP’s protection type.” App. 6 (Halabi) at pg. 189.

492. As another example, Halabi provides a detailed discussion of how a secondary Pseudowire that has been configured based on various Pseudowire protection configuration parameters can be used in a dual-homing scenario. *See, e.g., id.* at pgs. 101-103.

493. As another example, Halabi teaches various “GMPLS Protection and Mechanisms” that utilize a standby Pseudowire that has been configured based at least in part on the Pseudowire protection configuration parameter. *See, e.g., id.* at pgs. 188-189.

494. Each of these examples for Halabi discloses “using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter” under BSL’s apparent claim construction.

*f. determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

495. Claim 1 further recites: “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

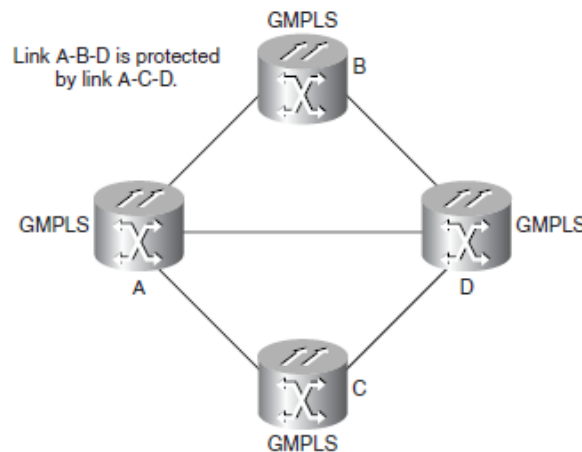
496. For example, Halabi teaches a “1+1 protection” mechanism where “[t]he data is transmitted simultaneously over the two disjoint paths. The receiver selects the working path . . . the backup paths may not be used by other LSPs because the data is transmitted on both paths.” *Id.* at pg. 189.

497. It is my understanding that BSL has taken the position that “existing traffic on the standby Pseudowire” refers to any “working traffic” and that this can include duplicative traffic that is the same as the protected traffic. Moreover, I understand that BSL has interpreted “preempting” such traffic based on a

“priority” to encompass a situation where the backup traffic in a 1+1 protection scheme is dropped during normal operation (because, according to BSL, the primary Pseudowire has “priority” over a secondary or backup Pseudowire). Under the Patent Owner’s broad interpretation of this claim, Halabi’s disclosure of a “1+1 protection” mechanism whereby the receiver selects which traffic to use would disclose this claim element.

498. Halabi also teaches the preemption of “existing traffic” that is not duplicative of the protected traffic that is being transmitted on the working path. For example, Halabi teaches various forms of “preemption” as discussed above are depicted in Figure 8.5, as follows:

**Figure 8-5** *Link Protection Types*



Link A-B-D is protected by link A-C-D. Link A-C-D is of type Extra Shared. The following protection scenarios can occur:

- **Link A-B-D is 1+1 protected**—Link A-C-D protects link A-B-D. Link A-C-D is not advertised and hence does not carry any LSPs unless link A-B-D fails.
- **Link A-B-D is 1:1 protected**—Link A-C-D protects link A-B-D. Link A-C-D is advertised and can carry LSPs, but it gets preempted to protect link A-B-D if link A-B-D fails.

App. 6 (Halabi) at pgs. 174-175.

499. Further to this example, Halabi teaches a “1:1 protection” mechanism where a dedicated backup path is preallocated to protect a primary path, and a “M:N protection” mechanism where M backup paths are preallocated to protect N



primary paths. *Id.* at pg. 189. In each of these schemes, Halabi teaches that “the backup paths may be used by other LSPs.” *Id.* It would be apparent to one skill in the art that, if the backup paths are used by other LSPs, and if the primary path fails, it would necessarily require the preemption of existing traffic on the backup path in order to achieve the protection of the primary path. Thus, under BSL’s apparent interpretation of “priority” (*i.e.*, as any “preference,” including the designation as primary vs. secondary) these disclosures teach preempting the existing traffic (*i.e.*, the traffic that is transferred on the backup LSP by other LSPs during normal operation), based, at least in part on a priority (*i.e.*, the designation as a “backup” instead of a “primary” path).

500. As another example, Halabi teaches that “[t]he resource allocated for a secondary LSP may be used by other LSPs until the primary LSP fails over to the secondary LSP. At that point, any set of LSPs that are using the resources for the secondary LSP must be preempted.” App. 6 (Halabi) at pg. 184 (“Protection Information”). This is another example of how Halabi discloses “preempting existing traffic” based (at least in part) on a priority (*e.g.*, the designation as secondary vs. primary) under BSL’s apparent interpretation of this element.

501. As another example, Halabi teaches that a “preemption attribute” can be used to “determine whether a traffic trunk can preempt another traffic trunk from a given path.” App. 6 (Halabi) at pg. 128, ¶ 5. Halabi further teaches that “[p]reemption can be used to ensure that high-priority traffic can always be routed in favor of lower-priority traffic that can be preempted. Service providers can use this attribute to offer varying levels of service. A service that has preemption could be priced at a higher rate than a regular service.” *Id.* It would be apparent to one skilled in the art that “routing” high-priority traffic in favor of lower-priority traffic would require preemption of the lower-priority traffic based on a priority that has been assigned to that traffic under BSL’s apparent claim construction.

502. As another example, Halabi discloses a “policing attribute” that “indicates whether traffic that does not conform to a certain SLA should be passed, rate limited, *dropped* or marked for further action.” *Id.* at pg. 129. It would be apparent to one skilled in the art that such a “policing” parameter would be used to make determinations about whether to preempt existing traffic, at least under BSL’s apparent claim construction.

503. Halabi also discloses a “resilience attribute” that “specif[ies] detailed actions to be taken under failure, such as the use of alternate paths, and specify the rules that govern the selection of these paths.” *Id.* at pg. 128. It would be apparent to one skilled in the art that such a “resilience” parameter would be used to make determinations about whether to preempt existing traffic, at least under BSL’s apparent claim construction.

504. As another example, the “QoS” in the “FLOW\_SPEC Object” can be used to preempt traffic that does not match any of the filter specs are preempted as “best-effort traffic,” at least under BSL’s apparent claim construction. *Id.* at pg. 145.

505. As another example, Halabi teaches that “[t]he SESSION\_ATTRIBUTE object allows RSVP-TE to set different LSP priorities, preemption and fast-reroute features” and these “are used to select alternate LSPs in case of a failure in the network.” *Id.* at pg. 144. Further to this example, Halabi teaches that the SESSION\_ATTRIBUTE can include “fields such as Setup Priority and Holding Priority, which affect whether this session can preempt or can be preempted by other sessions.” *Id.* These are additional examples of how “existing traffic” (*i.e.*, the traffic in a current session) could be preempted based on a “priority” (*i.e.*, the setup or holding priority) that has been assigned to the standby Pseudowire under BSL’s apparent claim construction.

506. In sum, it is my opinion that Halabi discloses each and every element of claim 1 of the '652 patent under BSL's apparent claim constructions, and thus Halabi anticipates claim 1.

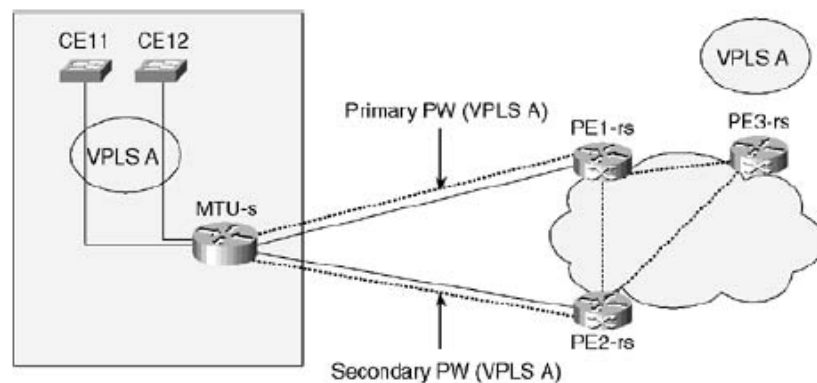
2. *Claim 2: A method as recited in claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.*

507. Claim 2 recites: "a method as recited in claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire."

508. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL's apparent claim construction, Halabi discloses the additional elements of claim 2.

509. For example, Halabi teaches that Dual-Homed MTU Device techniques can be used to set up a "primary" and a "secondary" Pseudowire whereby the "secondary" Pseudowire is configured to protect the "primary" Pseudowire as shown in Figure 4-22:

**Figure 4-22. Dual-Homed MTU Device**



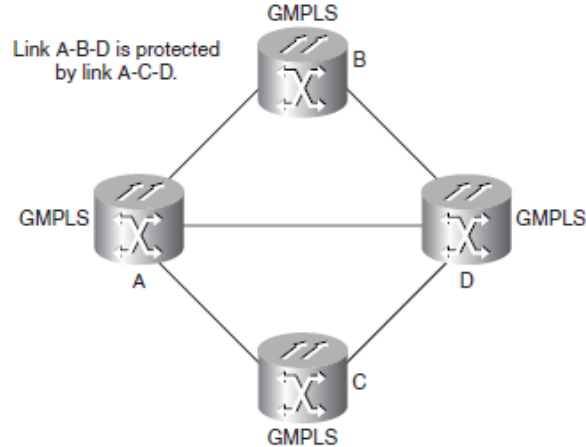
App. 6 (Halabi) at Fig. 4-22; *see also id.* at pgs. 102-103.

510. As another example, Halabi teaches that “Protection information also indicates whether the LSP is a primary or secondary LSP. A secondary LSP is a backup to a primary LSP. The resources allocated for a secondary LSP are not used until the primary LSP fails. The resources allocated for a secondary LSP may be used by other LSPs until the primary LSP fails over to the secondary LSP. At that point, any set of LSPs that are using the resources for the secondary LSP must be preempted.” *Id.* at pg. 184.

511. As another example, Halabi teaches that GMPLS signaling protocols can be used to configure a standby Pseudowire that is set up to provide protection to at least one primary Pseudowire in a variety of protection schemes. App. 6 (Halabi) at pg. 189. For example, Halabi teaches that in 1+1 protection, “[t]he data is transmitted simultaneously over the two disjoint paths.” *Id.* As another example, Halabi teaches that in 1:1 protection, “a dedicated **backup** path is preallocated to protect the **primary** path.” *Id.* As another example, Halabi teaches that in M:N protection, “M backup paths are preallocated to protect N primary paths. However, data is not replicated onto a **backup** path, but only transmitted in case of failure on the **primary** path.” *Id.*

512. As another example, Halabi depicts various protection schemes where a backup path is set up to protect a primary path in Figure 8.5:

Figure 8-5 Link Protection Types



Link A-B-D is protected by link A-C-D. Link A-C-D is of type Extra Shared. The following protection scenarios can occur:

- **Link A-B-D is 1+1 protected**—Link A-C-D protects link A-B-D. Link A-C-D is not advertised and hence does not carry any LSPs unless link A-B-D fails.
- **Link A-B-D is 1:1 protected**—Link A-C-D protects link A-B-D. Link A-C-D is advertised and can carry LSPs, but it gets preempted to protect link A-B-D if link A-B-D fails.

App. 6 (Halabi) at Fig. 8.5.

513. As another example, Halabi teaches various traffic trunk operations and attributes, such as the “preemption” and “protection” attributes that can be used to configure a standby PW to provide protection to at least one primary PW. *Id.* at pg. 128.

514. These are just some examples of the numerous disclosures in Halabi where a “standby” path is configured to protect at least one “primary” path under BSL’s apparent claim constructions.

515. In sum, it is my opinion that Halabi discloses each and every element of claim 2 of the ’652 patent under BSL’s apparent claim constructions, and thus Halabi anticipates claim 2.

3. *Claim 3: A method as recited in claim 1 wherein the standby Pseudowire is configured to provide protection to at least one*

*primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.*

516. Halabi discloses “a method as recited in claim 1 wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.”

517. I incorporate by reference the portions of this declaration pertaining to Claim 1 and 2 above. Under BSL’s apparent claim construction, Halabi discloses the additional elements of claim 3.

518. Halabi further teaches that “in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one primary Pseudowire to the standby Pseudowire.”

519. For example, Halabi teaches that, in a dual-homed MTU device, upon failure of the primary Pseudowire, “the MTU-s immediately switches to the secondary Pseudowire. At this point the PE2-rs that is terminating the secondary PW starts learning MAC addresses on the spoke PW.” App. 6 (Halabi) at pgs. 102-103.

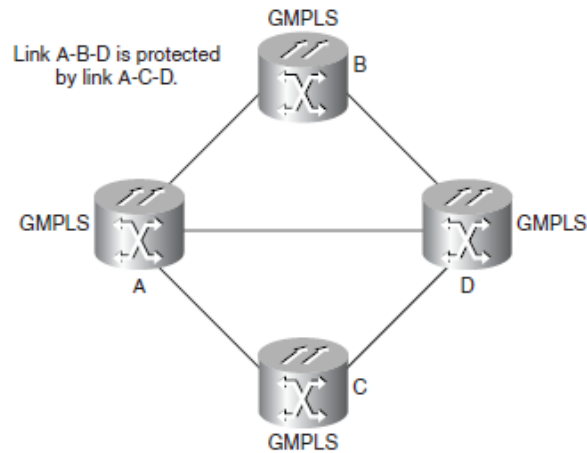
520. As another example, Halabi teaches that GMPLS signaling protocols can be used to configure a standby Pseudowire that is set up to provide protection to at least one primary Pseudowire in a 1+1, 1:1 or m:n protection scheme, whereby traffic is switched over from the primary to the backup when there is a failure on the primary Pseudowire. App. 6 (Halabi) at pg. 189.

521. As another example, Halabi teaches that “Protection information also indicates whether the LSP is a primary or secondary LSP. A secondary LSP is a backup to a primary LSP. The resources allocated for a secondary LSP are not used until the primary LSP fails. The resources allocated for a secondary LSP may

be used by other LSPs until the primary LSP fails over to the secondary LSP. At that point, any set of LSPs that are using the resources for the secondary LSP must be preempted.” *Id.* at pg. 184.

522. As another example, Halabi depicts various protection schemes where a backup path is set up to protect a primary path in Figure 8.5:

**Figure 8-5** *Link Protection Types*



Link A-B-D is protected by link A-C-D. Link A-C-D is of type Extra Shared. The following protection scenarios can occur:

- **Link A-B-D is 1+1 protected**—Link A-C-D protects link A-B-D. Link A-C-D is not advertised and hence does not carry any LSPs unless link A-B-D fails.
- **Link A-B-D is 1:1 protected**—Link A-C-D protects link A-B-D. Link A-C-D is advertised and can carry LSPs, but it gets preempted to protect link A-B-D if link A-B-D fails.

App. 6 (Halabi) at Fig. 8.5. In connection with this diagram, Halabi teaches that, when there is a failure, the LSPs switch to the pre-designated backup path. *Id.*

523. In sum, it is my opinion that Halabi discloses each and every element of claim 3 of the '652 patent under BSL's apparent claim constructions, and thus Halabi anticipates claim 3.

4. *Claim 4: A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.*

524. Claim 4 recites: “A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.”

525. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL’s apparent claim construction, Halabi discloses the additional elements of claim 4.

526. For example, Halabi teaches a “1+1 protection” where “[t]he data is transmitted simultaneously over the two disjoint paths. The receiver selects the working path based on the best signal.” App. 6 (Halabi) at pg. 189. Under BSL’s apparent construction, this “1+1” operation discloses the elements of claim 4.

527. As another example, Halabi teaches a “M:N protection” where “M backup paths are preallocated to protect N primary paths” and “the backup paths may be used by other LSPs.” App. 6 (Halabi) at pg. 189. It would be apparent to one skilled in the art that “M:N protection” would require a dynamic selection of the standby Pseudowire because  $N > M$  and other traffic is allowed to use the M connections while the N protected paths are operating normally. *Id.*

528. As another example, Halabi teaches that “Traffic Engineering” allows you to “dynamically build explicit LSPs.” *Id.* at pg. 125. For example, one can “caus[e] a trunk to reroute from its original path via manual or dynamic configuration.” *Id.* at pg. 127.

529. As another example, Halabi teaches that RSVP can be used to “dynamically reroute an established LSP tunnel.” *Id.* at pg. 137.

530. As another example, Halabi teaches that GMPLS can be used for “dynamic circuit provisioning” that “can be used to establish point-to-point or multipoint-to-point virtual private optical networks.” *Id.* at pg. 152.



531. Moreover, Halabi discloses that the decisions can be made about how to send traffic in real-time using the preemption attribute and/or priority attribute (App. 6 (Halabi) at pg. 128), which inherently requires the dynamic selection of a backup path upon failure under BSL's apparent claim construction.

532. In sum, it is my opinion that Halabi discloses each and every element of claim 4 of the '652 patent under BSL's apparent claim constructions, and thus Halabi anticipates claim 4.

5. *Claim 5: A method as recited in claim 1, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

533. Claim 5 recites: "A method as recited in claim 1, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme."

534. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL's apparent claim construction, Halabi discloses the additional elements of claim 5.

535. For example, as mentioned above, Halabi teaches that a "TLV" can be used to carry "protection information" and is thus an embodiment of a "Pseudowire protection configuration parameter" that includes a "protection property" under BSL's apparent construction. In connection with this example, Halabi teaches that "GMPLS uses a new object type length value (TLV) field to carry LSP protection information . . . Protection information indicates the LSP's **link protection type** . . ." App. 6 (Halabi) at pg. 184, ¶¶ 5-6. Halabi further teaches that a of different "link protection types":

## Link Protection Types

GMPLS introduces the concept of a *link protection type*, which indicates the protection capabilities that exist for a link. Path computation algorithms use this information to establish links with the appropriate protection characteristics. This information is organized in a hierarchy where typically the minimum acceptable protection is specified at path instantiation and a path selection technique is used to find a path that satisfies at least the minimum acceptable protection. The different link protection types are as follows:

- **Extra Traffic**— This type of link protects another link or links. In case of failure of the protected links, all LSPs on this link are lost.
- **Unprotected**— This type of link is simply not protected by any other link. If the unprotected link fails, all LSPs on the link are lost.
- **Shared**— This type of link is protected by one or more disjoint links of type Extra Traffic.
- **Dedicated 1:1**— This type of link is protected by a disjoint link of type Extra Traffic.
- **Dedicated 1+1**— This type of link is protected by a disjoint link of type Extra Traffic. However, the protecting link is not advertised in the link-state database and therefore is not used by any routing LSPs.
- **Enhanced**— This type of link indicates that a protection scheme that is more reliable than Dedicated 1+1 should be used—for example, four-fiber BLSR.

App. 6 (Halabi) at pg. 174.

536. These “link protection types” are similar to the “protection schemes” disclosed in the ’652 patent and thus, the inclusion of Halabi’s “link protection type” in the TLV message is an embodiment of “wherein the protection property further includes . . . a protection scheme” under BSL’s apparent constructions.

537. In sum, it is my opinion that Halabi discloses each and every element of claim 5 of the ’652 patent under BSL’s apparent claim constructions, and thus Halabi anticipates claim 5.

8. *Claim 8: A method as recited in claim 5, wherein the protection scheme indicates at least one of the following:*

538. Claim 8 recites: “A method as recited in claim 5, wherein the protection scheme indicates at least one of the following.”

539. I incorporate by reference the portions of this declaration pertaining to Claims 1 and 5 above.

a. *a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

540. Claim 8 further recites: “a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires. Under BSL’s apparent claim construction, Halabi discloses this element.

541. For example, Halabi discloses a 1+1 protection scheme:

GMPLS uses the following protection mechanisms:

- **1+1 protection**— The data is transmitted simultaneously over the two disjoint paths. The receiver selects the working path based on the best signal.

App. 6 (Halabi) at pg. 189; *see also id.* at pg. 174.

b. *a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

542. Claim 8 further recites: “a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

543. For example, Halabi discloses a 1:1 protection scheme:

- **1:1 protection**— A dedicated backup path is preallocated to protect the primary path.

App. 6 (Halabi) at pg. 189; *see also id.* at pg. 174.

c. *a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires;*

544. Claim 8 further recites: “a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires.”

545. Halabi does not expressly disclose a 1:N scheme. However, a 1:N protection scheme is a special kind of M:N protection scheme. As shown below, Halabi discloses an M:N protection scheme. As such, it is my opinion that this limitation would have been inherent and/or obvious in light of the disclosure in Halabi concerning M:N protection.

d. or an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.

546. Claim 8 further recites: “or an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.” Under BSL’s apparent claim construction, Halabi discloses this element.

547. For example, Halabi discloses a M:N protection scheme:

- **M:N protection**— M backup paths are preallocated to protect N primary paths. However, data is not replicated onto a backup path, but only transmitted in case of failure on the primary path.

For 1:1 and M:N protection, the backup paths may be used by other LSPs. For 1+1 protection, the backup paths may not be used by other LSPs because the data is transmitted on both paths.

App. 6 (Halabi) at pg. 189; *see also id.* at pg. 174.

548. In sum, it is my opinion that Halabi discloses each and every element of claim 8 of the ’652 patent under BSL’s apparent claim constructions, and thus Halabi anticipates claim 8.

**Claims 9-11, 13-15 and 17**

549. I note that the limitations of independent claim 9 and claim 14 are nearly identical to claim 1, except for the fact that claim 9 is a system claim and claim 14 is a computer program claim. The dependent claims also mirror the claims that depend on claim 1. For example, claim 10 is analogous to claim 2, claims 11 and 15 are analogous to claim 5, and claims 13 and 17 are analogous to claim 8. As such, my analysis below largely incorporates by reference my analysis with respect to claims 1-5 and 8. The claims correlate as follows:

<u>Method</u>	<u>System</u>	<u>Computer Program</u>
Claim 1	Claim 9	Claim 14
Claim 2	Claim 10	
Claim 3		
Claim 4		

Claim 5	Claim 11	Claim 15
Claim 8	Claim 13	Claim 17

9. *Claim 9: A system for providing protection to network traffic, comprising a processor configured to:*

550. Claim 9 recites: “a system for providing protection to network traffic.” Under BSL’s apparent claim construction, Halabi discloses this element.

551. I incorporate by reference my comments from claim 1 above.

552. One skilled in the art would understand that the protection system described by Halabi requires the use of various computer hardware, including routers and/or other network devices that contain processors.

*a. send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;*

553. Claim 9 further recites: “send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

554. I incorporate by reference my comments from claim element 1(a) and 1(b) above.

- b. receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

555. Claim 9 further recites: “receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, Halabi discloses this element.

556. I incorporate by reference my comments from claim element 1(c) above.

- c. accept the Pseudowire protection configuration parameter by the destination node;*

557. Claim 9 further recites: “accept the Pseudowire protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, Halabi discloses this element.

558. I incorporate by reference my comments from claim element 1(d) above.

- d. use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

559. Claim 9 further recites: “use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter.” Under BSL’s apparent claim construction, Halabi discloses this element.

560. I incorporate by reference my comments from claim element 1(e) above.

- e. *determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

561. Claim 9 further recites: “determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

562. I incorporate by reference my comments from claim element 1(f) above.

563. In sum, it is my opinion that Halabi discloses each and every element of claim 9 of the ’652 patent under BSL’s apparent claim constructions, and thus Halabi anticipates claim 9.

10. *Claim 10: A system as recited in claim 9, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.*

564. Claim 10 recites: “A system as recited in claim 9, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.”

565. I incorporate by reference my comments from claim 1, claim 2 and claim 9 above. Under BSL’s apparent claim construction, Halabi discloses the additional elements of this claim.

566. In sum, it is my opinion that Halabi discloses each and every element of claim 10 of the ’652 patent under BSL’s apparent claim constructions, and thus Halabi anticipates claim 10.

11. *Claim 11: A system as recited in claim 9, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

567. Claim 11 recites: “A system as recited in claim 9, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.”

568. I incorporate by reference my comments from claim 1, 5 and 9 above. Under BSL’s apparent claim construction, Halabi discloses the additional elements of this claim.

569. In sum, it is my opinion that Halabi discloses each and every element of claim 11 of the ’652 patent under BSL’s apparent claim constructions, and thus Halabi anticipates claim 11.

13. *Claim 13: A system as recited in claim 11, wherein the protection scheme indicates at least one of the following:*

570. Claim 13 recites: “A system as recited in claim 11, wherein the protection scheme indicates at least one of the following.”

571. I incorporate by reference my comments from claims 1, 8, 9, and 11 above. Under BSL’s apparent claim construction, Halabi discloses the additional elements of this claim.

a. *a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

572. Claim 13 further recites: “a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires.” Under BSL’s apparent claim construction, Halabi discloses this element.

573. I incorporate by reference my comments from claim 8 above. Under BSL’s apparent claim construction, Halabi discloses this element.



*b. a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

574. Claim 13 further recites: “a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

575. I incorporate by reference my comments from claim 8 above.

*c. an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.*

576. Claim 13 further recites: “an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.” Under BSL’s apparent claim construction, Halabi discloses this element.

577. I incorporate by reference my comments from claim 8 above.

578. In sum, it is my opinion that Halabi discloses each and every element of claim 13 of the ’652 patent under BSL’s apparent claim constructions, and thus Halabi anticipates claim 13.

*14. Claim 14: A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for:*

579. Claim 14 recites: “a computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for.” Under BSL’s apparent claim construction, Halabi discloses this element.

580. I incorporate by reference my comments from claim 1 and claim 9 above.

- a. *sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;*

581. Claim 14 recites: “sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

582. I incorporate by reference my comments from claim element 1(a) and 1(b) above.

- b. *receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

583. Claim 14 recites: “receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, Halabi discloses this element.

584. I incorporate by reference my comments from claim element 1(c) above.

- c. accept the Pseudowire protection configuration parameter by the destination node;*

585. Claim 14 recites: “accept the Pseudowire protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, Halabi discloses this element.

586. I incorporate by reference my comments from claim element 1(d) above.

- d. using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

587. Claim 14 recites: “using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter.” Under BSL’s apparent claim construction, Halabi discloses this element.

588. I incorporate by reference my comments from claim element 1(e) above.

- e. determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

589. Claim 14 recites: “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

590. I incorporate by reference my comments from claim element 1(f) above.

591. In sum, it is my opinion that Halabi discloses each and every element of claim 14 of the '652 patent under BSL's apparent claim constructions, and thus Halabi anticipates claim 14.

15. *Claim 15: A computer program product as recited in claim 14, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

592. Claim 15 recites: "A computer program product as recited in claim 14, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme." Under BSL's apparent claim construction, Halabi discloses this element.

593. I incorporate by reference my comments from claim 1, 5 and 14 above.

594. In sum, it is my opinion that Halabi discloses each and every element of claim 15 of the '652 patent under BSL's apparent claim constructions, and thus Halabi anticipates claim 15.

17. *Claim 17: A computer product as recited in claim 15, wherein the protection scheme indicates at least one of the following:*

595. Claim 17 recites: "a computer product as recited in claim 15, wherein the protection scheme indicates at least one of the following." Under BSL's apparent claim construction, Halabi discloses this element.

596. I incorporate by reference my comments from claim 1, 5, 8, 14, and 15 above.

a. *a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

597. Claim 15 further recites: "a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires." Under BSL's apparent claim construction, Halabi discloses this element.

598. I incorporate by reference my comments from claim 8 above.

*b. a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

599. Claim 15 further recites: “a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire.” Under BSL’s apparent claim construction, Halabi discloses this element.

600. I incorporate by reference my comments from claim 8 above.

*c. an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.*

601. Claim 15 further recites: “an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.” Under BSL’s apparent claim construction, Halabi discloses this element.

602. I incorporate by reference my comments from claim 8 above.

603. In sum, it is my opinion that Halabi discloses each and every element of claim 17 of the ’652 patent under BSL’s apparent claim constructions, and thus Halabi anticipates claim 17.

604. Thus, it is further my opinion that Halabi anticipates each of the Challenged Claims of the ’652 patent.

**G. Halabi renders Claims 1-5, 8-11, 13-15, and 17 obvious under 35 U.S.C. § 103.**

605. If certain aspects recited in claims 1-5, 8-11, 13-15, and 17 are not deemed to be disclosed or inherent over Halabi, then it is my opinion that the inclusion of those aspects certainly would be obvious over Halabi.

606. For example, should Halabi be found to disclose each of specific claim limitations (a) – (f) of the independent claims in the context of traditional LSP tunnels, trunks, and/or GMPLS paths, as opposed to Pseudowires, claims 1-5, 8-11, 13-15, and 17 are nevertheless obvious because applying these techniques to Pseudowires would have been nothing more than a “predictable variation.”

607. Indeed, Halabi itself discusses the concept of Pseudowire and Pseudowire protection in detail (*see, e.g.*, Ch. 4).

608. Moreover, the Examiner found in the original prosecution of the '652 patent that, at the time of the invention, it would have been obvious to a person of skill in the art to apply the protection/ restoration mechanisms used in the context of MPLS and/or GMPLS to a pseudowire environment because both MPLS/GMPLS and Pseudowire are in the narrow field of point-to-point virtual links. I agree with the Examiner that this would be an obvious and predictable combination.

609. I also note that the Patent Owner did not contest that combining MPLS/GMPLS protection/restoration mechanisms with Pseudowire and Pseudowire protection concepts was obvious when traversing the Examiner's rejection of the claims that ultimately issued as claims 1, 9 and 14.

610. Moreover, the '652 patent itself admits that it is obvious to apply protection schemes from other network environments, such as MPLS, to Pseudowires in that it describes the existing protection methods for Pseudowires as including “MPLS Fast Reroute.” App. 2 ('652 patent) at pg. 1:49-64.

611. The close connection between MPLS and Pseudowire is further evidenced by the fact that concepts pertaining to MPLS/GMPLS and Pseudowire are all described in the same *Metro Ethernet* book. Moreover, the whole point of Halabi is to discuss concepts, protocols, and traffic engineering techniques that will allow consistency and resiliency in hybrid Layer 2 and Layer 3 networks that use Pseudowire, which Halabi teaches are necessary to deploy Ethernet in the Metro. Moreover, these topics are frequently discussed together. *See, e.g.*, App. 22 (L2 VPN Architectures).

612. And, Halabi itself notes that “[w]hen traffic moves from one site to another across the carrier’s backbone [via PW], it follows the MPLS label switched path (LSP) assigned for that traffic . . . ***the LSP could be traffic-engineered*** . . . many mechanisms can be used for traffic rerouting in case of failure.” App. 6 (Halabi) at pg. 80. Thus, Halabi itself expressly contemplates that the disclosed MPLS and GMPLS traffic engineering techniques (**including preemption and priority**) can be used in connection with Pseudowire protection, not just LSP tunnel protection. *Id.*

613. In sum, it is my opinion that Halabi alone renders each of the Challenged Claims obvious under 35 U.S.C. § 103.

**H. Halabi in view of RFC 3386 and Owens renders Claims 1-5, 8-11, 13-15, and 17 obvious under 35 U.S.C. § 103.**

614. If certain aspects recited in claims 1-5, 8-11, 13-15, and 17 are not deemed to be disclosed, inherent or obvious over Halabi, then it is my opinion that the inclusion of those aspects certainly would be obvious over Halabi in view of RFC 3386 and Owens.

615. Detailed discussions of RFC 3386 and Owens can be found in Grounds 3 and 4, which I hereby incorporate by reference.

616. As described above, Halabi is a book that provides a general summary of the state of the art concerning MPLS, PW, GMPLS, and traffic engineering as it applies to deploying Ethernet in the Metro. App. 6 (Halabi) at xv.

617. Many of the concepts in Halabi are summaries of IETF industry standards documents. *See, e.g.*, App. 6 (Halabi) at vi (thanking “many of the authors of the IETF RFCs and IETF drafts whose information has been used for some of the concepts and definitions in this book”). Indeed, Halabi summarizes several RFCs and Internet Drafts regarding MPLS, traffic engineering and Pseudowires that I have discussed above, including draft-martini, RSVP-TE protocol (RFC 3209) and the Lasserre Draft. *See, e.g.*, App. 6 (Halabi) at Chapter 4 (draft-martini and Lasserre Draft) and Chapter 6 (RFC 3209). I further note that a quick review of Halabi demonstrates that there are numerous references to various IETF RFCs and Internet Drafts.

618. As also mentioned above, RFC 3386 is an IETF standard that describes various configuration parameters that can be used to provide traffic protection in a wide range of networks, including the MPLS, GMPLS, and Pseudowire environments discussed in Halabi. App. 5 (RFC 3386).

619. As described above, Owens is a patent invented by members of the IETF that discloses a specific method for applying these protection techniques in



an MPLS network. It covers the range of configuration options, and also provides a detailed explanation regarding how to configure those options. App. 15 (Owens) at 6:56-59, 7:1-6; *see also id.* at 5:23-29; 1:34-36.

620. The protection techniques described in RFC 3386 and Owens were widely-known at the time of Halabi. Given the prominence of these methods, and their close relationship to the protocols and concepts that Halabi summarizes, it would have been obvious to apply the well-known protection techniques and parameters described in RFC 3386 and Owens to the well-known Pseudowire environment discussed by Halabi.

621. A motivation to combine Halabi with RFC 3386 and Owens also exists. Indeed, the references are in the same very narrow field of L2-L3 internet protocols and are directed to the same problem (obtaining greater control over configuration, management and resiliency of network environments).

622. In fact, Halabi itself teaches that the whole point of employing Pseudowire technology is to allow native Layer 2 services to take advantage of the scalability and **reliability** mechanisms of MPLS: “hybrid Layer 2 (L2) and Layer 3 (L3) IP and MPLS networks [*i.e.*, PW networks] have emerged as a solution that marries Ethernet’s simplicity and cost effectiveness” with the “scalability and **reliability** “ that “exist only in IP and Multiprotocol Label Switching (MPLS) control planes.” App. 6 (Halabi) at xv. Thus, combining Halabi with RFC 3386 and Owens is not only obvious, but also specifically encouraged by Halabi.

#### **Claims 1, 9, 14**

623. If certain aspects recited in claims 1, 9 and 14 are not deemed to be disclosed, inherent or obvious over Halabi alone, it is my opinion that the inclusion of those aspects certainly would be obvious over Halabi in view of RFC 3386 and Owens.

**“determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire”**

624. Claims 1, 9 and 14 recite in part: “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.”

625. To the extent that Halabi is deemed to not expressly or inherently disclose this element, it is my opinion that this element would have been obvious in view of RFC 3386 and Owens.

626. As noted in Grounds 5 and 6, Halabi discloses “preempting existing traffic” based on “the priority for the standby Pseudowire” in the context of a 1+1 protection scheme. Under BSL’s interpretation of this element—which encompasses dropping duplicative traffic on a standby path during normal operation (*i.e.*, BSL’s version of “preempting existing traffic”) based on the fact that it is a *standby* path (as opposed to protected path) (*i.e.*, BSL’s version of “the priority”)—Halabi clearly discloses this element.

627. Halabi also discloses other types of traffic preemption. App. 6 (Halabi) at pgs. 128 (“preemption attribute”); 189 (1:1 and M:N where “backup paths may be used by other LSPs”).

628. RFC 3386 and Owens contain more specific disclosures of these same types of preemption based on priority. For example, RFC 3386 discloses “preempting existing traffic” based on the “relative priority” assigned to the protection Pseudowire. App. 5 (RFC 3386) at § 2.2.2 (“Extra traffic, also referred to as preemptable traffic, is the traffic carried over the protection entity while the working entity is active. Extra traffic is not protected, *i.e.*, **when the protection entity is required to protect the traffic that is being carried over the working entity, the extra traffic is preempted.**”); § 2.3 (“In the 1:n protection architecture

. . . [w]hen multiple working entities have failed simultaneously, only one of them can be restored by the common protection entity. This contention could be resolved by *assigning a different preemptive priority* to each working entity.”).

629. As another example, Owens teaches that, in a 1:1 protection scheme, “the working traffic normally travels only on the working path, and is switched to the protection path only when the working entity is unavailable. *Once the protection switch is initiated, all the low priority traffic being carried on the protection path is discarded to free resources for the working traffic.*” App. 15 (Owens) at 7:1-6; *see also* 5:23-29; 1:34-36 (“[A] *protection priority* could be used as a differentiating mechanism for premium services.”).

630. Because Halabi already generally teaches the concepts of preemption of existing traffic (*see, e.g.*, App. 6 (Halabi) at pg. 128 (discussing preemption attribute); 175 (noting that in 1:1 protection, the protection link “gets preempted to protect [the primary link if it fails]”), as well as assigning relative priorities (*id.* at pg. 128 (discussing priority attribute); 144-145 (discussing Setup/Holding Priorities)), it would have been an obvious and predictable step to use those priorities to make decisions about preemption during a network failure and to preempt existing traffic on the standby path, as taught by RFC 3386 and Owens.

**I. Chen in view of Voit and Blanchet renders Claims 1-5, 8-11, 13-15, and 17 obvious under 35 U.S.C. § 103.**

631. “The LSP Protection/Restoration Mechanism in GMPLS” is a paper by Ziyang Chen from the University of Ottawa that is dated October 1, 2002 (“Chen”).

632. Chen is a report that discusses GMPLS and specifies how different protocols can contribute to path protection and restoration in GMPLS. The report illustrates how to signal these protection mechanisms, and also illustrates how they work. *See* App. 7 (Chen) at pg. 3.

633. Chen was cited by the Examiner in several office actions as anticipating the ’652 patent in view of Voit and Blanchet.

634. As I described above, the applicant distinguished Chen on the grounds that it supposedly does not disclose the element of “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” More specifically, the applicant argued that “Chen describes a type of link protection in which **backup lines will not transport traffic** . . . .” App. 3 (’652 File History) at 082. In support of this argument, the applicant cited a diagram in Chen of “dedicated 1+1 link protection” as showing that “traffic is switched over from the primary link to the backup link when the primary link fails.” *Id.* at 084. Then, the applicant concluded that “[s]ince Chen describes the backup LSP as not transporting traffic until the primary LSP fails . . . Chen does not describe or even suggest ‘determining whether to **preempt existing traffic** on the standby Pseudowire”) *Id.* (emphasis in original).

635. Based on my review of BSL’s infringement contentions in the Concurrent Litigation, it is clear that BSL is now contending that the Challenged Claims can encompass the 1+1 protection scheme that the applicant disclaimed

during prosecution. For example, BSL's infringement contentions for the element of "determining whether to preempt existing traffic on the standby Pseudowire" are as follows:

"When both the primary pseudowire and the standby pseudowire are operational, the remote PE router of the primary pseudowire and the remote PE router of the standby pseudowire both send traffic, over their respective pseudowires, to the local PE router.

The local PE router accepts traffic from the primary pseudowire and drops traffic from the standby pseudowire. When the primary pseudowire fails, the local PE router then accepts traffic from the standby pseudowire. **When the local PE router accepts traffic from the primary pseudowire and drops traffic from the standby pseudowire without the possibility of being interrupted by traffic from the standby pseudowire, traffic from the primary pseudowire preempts traffic from the standby pseudowire.**

The local PE router determining whether to drop traffic from the standby pseudowire is an embodiment of determining whether to preempt traffic on the standby Pseudowire."

App. 25 (BSL's Preliminary Infringement Contentions) at '652 Chart, page 8 (bold added, underline in original). Thus, BSL is accusing a "1+1 protection scheme" where the same traffic is sent over both the primary and standby Pseudowire. BSL's theory is that the "existing traffic" is the duplicative traffic and that the "preemption" occurs when the local router drops traffic from the standby Pseudowire during normal operation. *Id.* This is precisely the protection scheme that the applicant distinguished during prosecution of the '652 patent.

636. To the extent that the Challenged Claims are construed as encompassing the 1+1 protection scheme described in Chen (and in Saleh), it is my opinion that the Challenged Claims should not have issued over Chen in view of Voit and Blanchet because they are obvious under 35 U.S.C. § 103.

637. As the Examiner found during prosecution of the '652 patent, Chen discloses a method of protecting traffic in a GMPLS network that involves setting up a backup path by sending a configuration parameter that includes a protection property between a source node and a destination node. Moreover, Chen also discloses that the protection property can include a priority and that the priority can be used to determine whether to preempt traffic. As the Examiner also found during prosecution, while Chen does not specifically disclose this technique in a Pseudowire environment, Pseudowires and Pseudowire protection using a redundant Pseudowire were well-known at the time (as disclosed, for example, in Voit). The Examiner found during original prosecution that Chen did not disclose an "acknowledgement" as required by the claims. As discussed in detail below, I disagree with the Examiner, as Chen discloses the details of RSVP-TE signaling protocol (including the use of a RESV message that is sent from the destination node to the source node in response to a request to set up a path according to specified parameters in a GMPLS network) in a portion of Chen that the Examiner does not appear to have examined. In any event, "configuration acknowledgement" messages were well known in the art at the time (as disclosed, for example, in Blanchet) and it would have been obvious to include this technique in the Chen system because it was a part of the standard protocols used to signal LSPs, Pseudowires, and other virtual paths.

1. *Claim 1: A method of providing protection to network traffic, comprising,*

638. Claim 1 recites: “A method of providing protection to network traffic, comprising.” Under BSL’s apparent claim construction, Chen discloses this element.

639. For example, Chen states that “[t]he main emphasis of this report is on the path protection/restoration mechanisms that can be used with GMPLS.” App. 7 (Chen) at pg. 6.

a. *sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire,*

640. Claim 1 further recites: “sending a ...protection configuration parameter for configuring a standby ... between a source node and a destination node, the ... protection configuration parameter indicating a protection property associated with the standby ....” Under BSL’s apparent claim construction, Chen renders this element obvious.

641. As noted by the Examiner, Chen teaches this element. For example, Chen teaches that “label distribution protocols may carry the link protection type.” *Id.* at pg. 21, ¶ 4.

642. As further noted by the Examiner, Chen teaches path protection/restoration in the context of GMPLS, but does not expressly teach it in the Pseudowire and Pseudowire protection context. However, as further noted by the Examiner, Voit teaches Pseudowire (App. 20 (Voit) at [0011]) and also teaches

Pseudowire protection in the form of a “network topology [that] is provided with redundant pseudowire connections . . . “ (App. 20 (Voit) at [0046]).

643. As further noted by the Examiner during prosecution, at the time of the invention, it would have been obvious to a person of ordinary skill in the art to apply the protection/restoration mechanism disclosed by Chen to a Pseudowire environment because both MPLS and Pseudowires are point-to-point virtual links. App. 3 ('652 File History) at pg. 106. The Examiner also noted that Chen and Voit are in the same field (network transfer) and are directed to the same problem sought to be solved (data traffic protection), and that the combination would merely require the application of a known technique to a similar system to improve its reliability. *Id.*

644. In fact, the '652 patent itself admits that it is obvious to apply MPLS protection schemes to Pseudowires in that it describes the existing protection methods for Pseudowires as including “MPLS Fast Reroute.” App. 2 ('652 patent) at 1:49-64.

645. Notably, when traversing the Examiner’s rejection of the claims that ultimately issued as claims 1, 9 and 14, the applicant **did not contest** that combining the protection/restoration mechanism of Chen with the Pseudowire and Pseudowire protection concepts discussed in Voit was obvious.

646. For all of the reasons stated above, I agree with the Examiner that Chen discloses “sending a... protection configuration parameter for configuring a standby... between a source node and a destination node, the ... protection configuration parameter indicating a protection property associated with the standby....” and that it would have been obvious to one of skill in the art to apply the protection methods discussion in Chen to a Pseudowire environment to accomplish Pseudowire protection.



*b. the protection property including a priority for the standby Pseudowire*

647. Claim 1 further recites: “the protection property including a priority for the standby . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

648. For example, the Examiner found that Chen teaches that “the resource allocation has priorities (carried by the signaling protocol).” App. 7 (Chen) at pg. 21, ¶ 5. As another example, Chen teaches that “[t]he GMPLS signaling protocol carries a flag that indicates if the LSP being set up is primary or secondary.” *Id.*

649. In addition to the examples cited by the Examiner, Chen also teaches this element in its more detailed discussion of RSVP-TE and CR-LDP signaling protocols. More specifically, Chen teaches that in RSVP-TE signaling, a PATH Message is used to carry configuration information from the source host to the destination host, and that the PATH Message can include a SESSION\_ATTRIBUTE object that specifies the setup and hold priorities and local protection properties for the connection. App. 7 (Chen) at pg. 37, ¶ 3, lines 1-5; ¶ 4, lines 1-4; page 38, ¶ 2, lines 4-7. The SESSION\_ATTRIBUTE object allows RSVP-TE to set different LSP priorities that are used to select alternate LSP and preempt the existing traffic on this LSP.

650. As another example, with respect to CR-LDP signaling, Chen teaches that “the TLV structure [is used] to encode messages” (*id.* at pg. 48, ¶ 6, line 1) and that “CR-LDP defines a new set of TLV structures to support explicit routed signaling, traffic parameters, LSP set-up/holding priority, etc.” (*id.* at pg. 49, ¶ 2, lines 2-4). Chen further teaches that “R1 sends out the CR-LDP Label Request message carrying the constraint-based route TLV . . . The Label Request message may carry the Traffic Parameter TLV, which specifies the traffic parameters to be

sent” (*id.* at pg. 49, ¶ 5) and that “CR-LDP defines the Protection TLV, which includes: (1) link protection type; (2) indication of whether the path is primary or backup” (*id.* at pg. 51, ¶ 2, lines 2-4).

651. Thus, I agree with the Examiner that Chen discloses “the protection property including a priority for the standby . . . .”

*c. receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

652. Claim 1 further recites: “receiving a . . . configuration acknowledgement indicating whether the . . . protection configuration parameter has been accepted.” Under BSL’s apparent claim construction, Chen renders this element obvious.

653. During the original prosecution, the Examiner found that Chen fails to expressly disclose a “configuration acknowledgement.” Based on my review of the entire Chen reference, I disagree with the Examiner and it is my opinion that Chen does in fact disclose a “configuration acknowledgement” that indicates with the “protection configuration parameter has been accepted by the destination node.”

654. For example, Chen teaches that in RSVP-TE signaling, the destination node “sends back a RESV message back toward the sender . . . The RESV message communicates with every router to make a resource reservation.” App. 7 (Chen) at pg. 37, ¶ 4, lines 6-8. Chen further teaches that the RESV message contains a “LABEL object” that is used by each node along the path to reserve resources and that “[s]uch a label distribution procedure repeats until the Resv message arrives at the sender node. The LSP establishment is done.” *Id.* at pg. 38, ¶ 2, lines 7-16. Moreover, Chen teaches that “link protection type in the protection information

[that is carried by the label distribution protocols in RSVP-TE or LDP signaling] is one of the TE requirements (or a constraint for a LSP to be set up). So the LSP set-up will not continue if the desired link protection cannot be provided” *Id.* at pg. 41 ¶ 3. It is my opinion that the “RESV message” disclosed by Chen that is sent from the destination node back to the source node is an embodiment of “receiving a ... configuration acknowledgement indicating whether the ... protection configuration parameter has been accepted.”

655. Similarly, in its detailed discussion of CR-LDP signaling, Chen teaches that “R5, as the ending node of the LSP. . . responds [to the CR-LDP Label Request message that carries constraint-based route TLVs] with a CR-LDP Label Mapping message, which carries a Label TLV.” *Id.* at pg. 49, ¶ 5, line 1 – pg. 50, ¶ 1, line 1.

656. Thus, it is my opinion that Chen discloses “receiving a... configuration acknowledgement indicating whether the... protection configuration parameter has been accepted.”

657. Even if the original Examiner were correct, however, and Chen does not expressly teach this element, I agree with the Examiner that it would be obvious in view of Blanchet, which teaches the use of a “configuration acknowledgement.” App. 21 (Blanchet) at [0035].

658. Indeed, as noted by the Examiner, “it would have been obvious to a person of ordinary skill in the art to modify the system to send an ACK indicating the acceptance of the configuration parameters in the system disclosed by Chen in view of Voit in order to make the system more reliable. Both Chen in view of Voit and Blanchet are in the same field of endeavor (Network transfer).” App. 3 (’652 File History) at pgs. 105-106. In addition to the Examiner notes, sending ACK is very common practice in network communications. In fact, since communications

is done between two systems, it is hard to find a protocol that does not send acknowledgments.

659. As I described in above in connection with Section VII(E) (RFC 3386 in combination with RFC 3209), the use of a configuration acknowledgement message was a part of the other standard signaling protocols that were used to establish LSP tunnels and Pseudowires, such as LDP, draft-Martini, and even the Hofmeister method described above. Thus, a person of skill in the art would have found it obvious to implement a configuration acknowledgement within the context of the protection techniques disclosed in Chen because a “configuration acknowledgement” was a standard part of MPLS and Pseudowire signaling techniques.

660. Moreover, the Patent Owner did not contest that combining an acknowledgement message with a method for configuring a standby path was obvious when traversing the Examiner’s rejection of the claims that ultimately issued as claims 1, 9 and 14. Instead, the Patent Owner only argued the alleged novelty of determining whether to preempt existing traffic on the standby Pseudowire based, at least in part, on the priority for the standby Pseudowire. *See* App. 3 (File History) at pgs. 082-084.

*d. accepting the Pseudowire protection configuration parameter by the destination node;*

661. Claim 1 further recites: “accepting the . . . protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, Chen renders this element obvious.

662. For example, Chen teaches that “[t]he link protection type in the protection information is one of the TE requirements (or a constraint) for a LSP to be set up. So the LSP set-up will not continue if the desired link protection cannot be provided. App. 7 (Chen) at pg. 41.

663. Thus, I agree with the Examiner that Chen discloses “accepting the . . . protection configuration parameter by the destination node.”

*e. using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

664. Claim 1 further recites: “using the standby . . . that is configured based at least in part on the . . . protection configuration parameter.” Under BSL’s apparent claim construction, Chen renders this element obvious.

665. For example, on pages 52-55, Chen describes the use of 1+1, M:N, 1:N and 1:1 protection mechanisms, as well as how priorities are utilized in deciding which paths to protect/preempt.

666. Thus, I agree with the Examiner that Chen discloses “using the standby . . . that is configured based at least in part on the . . . protection configuration parameter.”

*f. determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

667. Claim 1 further recites: “determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.” Under BSL’s apparent claim construction, Chen renders this element obvious.

668. As noted above, Chen teaches that in RSVP-TE signaling, a PATH Message is used to carry configuration information from the source host to the destination host, and that the PATH Message can include a SESSION\_ATTRIBUTE object that specifies the setup and hold priorities and local protection properties for the connection. App. 7 (Chen) at pg. 37, ¶ 3, lines 1-

5; ¶ 4, lines 1-4; page 38, ¶ 2, lines 4-7. The SESSION\_ATTRIBUTE object allows RSVP-TE to set different LSP priorities that are used to select alternate LSP and preempt the existing traffic. *See also* Section VII(E) above.

669. As noted above, during the original prosecution, the Examiner originally found that Chen discloses this element, but the applicant distinguished Chen based on two grounds, neither of which is valid under the Patent Owner's broad constructions of the terms "existing traffic" and "priority."

670. More specifically, the Patent Owner argued that "Chen describes a type of link protection in which backup links will not transport traffic" and that in Chen's disclosure of dedicated 1+1 link protection, "traffic is switched over from the primary link to the backup link when the primary link fails." App. 3 (File History) at pgs. 082-084 (citing diagram showing that Chen's protection scheme utilized a dedicated 1+1 link protection). The Patent Owner also argued that the "priority" identified by the Examiner applied to resource allocations, not the backup LSP and that "Chen discloses that the resources allocated to the backup LSP may be used by other LSPs that have lower priorities until the primary fails," at which time, "all the other LSPs using the resource allocated for the backup LSP must be preempted." *Id.* at pg. 3. Thus, the Patent Owner argued, "Chen merely describes **preempting the use** of these **prioritized resources** by the other LSPs." *Id.* (emphasis in original).

671. As noted above, the Patent Owner has now proposed that "priority" be interpreted to include a mere designation as standby vs. backup, and that "existing traffic on the standby Pseudowire" be interpreted as any "working traffic transmitted on the standby path." The Patent Owner's infringement contentions make clear that it is interpreting "working traffic" to include traffic that is duplicative of the traffic that being sent on the primary path (and which is only used in case of a failure), such as in a 1+1 protection scheme. Moreover, under the

Patent Owner’s proposed constructions, “determining whether to preempt existing traffic on the standby . . . , wherein the determination is based, at least in part, on the priority for the standby . . . ,” includes the act of a source node using traffic from a primary path that is protected under a 1+1 protection mechanism (where duplicative traffic is simultaneously sent on a backup path) during normal operation (because the source node drops the traffic on the duplicative path based on the fact that the path is designated as a backup, which is a lower priority than a primary designation). App. 25 (BSL’s Preliminary Infringement Contentions) at 8 (“When the local PE router accepts traffic from the primary pseudowire and drops traffic from the standby pseudowire without the possibility of being interrupted by traffic from the standby pseudowire, traffic from the primary pseudowire preempts traffic from the standby pseudowire.”).

672. Under these constructions that are proposed by the Patent Owner, Chen anticipates this element because Chen teaches a 1+1 protection scheme whereby “[t]he same user data is transmitted simultaneously over the two paths. . .” App. 7 (Chen) at page 53, ¶ 1, lines 3-7; *see also* page 56, ¶ 2 (“the node. . . will copy the traffic and insert it into both links”). In other words, the “same user data” that is “transmitted simultaneously” is the “existing traffic.” Chen further teaches that, “after initialization, the receiver *takes the traffic from the primary* link. *When the primary link fails*, LMP Fault Management . . . is used to localize the failure . . . [and] *Node B simply selects the traffic from the backup link.*”). Thus, under BSL’s broad proposed constructions, “tak[ing] the traffic from the primary link” during normal operation would comprise “preempting existing traffic on the standby” in that the “same user data” that is being transmitted on the backup link is dropped by the receiver. And, under BSL’s broad proposed constructions, the determination is “based, at least in part, on the priority for the standby” in that the selection of traffic by the receiver is based on

whether the link is designated as “primary” or “backup” (wherein “primary” has a higher “priority” than “backup”).

673. In sum, it is my opinion that Chen in view of Voit and Blanchet renders claim 1 of the ’652 patent under Patent Owner’s proposed interpretation of the claims.

2. *Claim 2: A method as recited in claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.*

674. Claim 2 recites: “A method as recited in claim 1, wherein the standby . . . is configured to provide protection to at least one primary . . . .”

675. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL’s apparent claim construction, Chen renders this claim obvious.

676. For example, Chen teaches that “[t]here are two LSP roles: **primary or secondary (backup)**. The GMPLS signaling protocol carries a flag that indicates . . . the resource allocated for a backup LSP may be used by an LSP that has lower priority until the primary LSP fails and the traffic is switched over to the backup.” App. 7 (Chen) at pg. 21, last paragraph.

677. In sum, it is my opinion that Chen renders claim 2 of the ’652 patent obvious in view of Voit and Blanchet under the Patent Owner’s proposed constructions of “priority” and “existing traffic.”



3. *Claim 3: A method as recited in claim 1 wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.*

678. Claim 3 recites: “A method as recited in claim 1 wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.”

679. I incorporate by reference the portions of this declaration pertaining to Claims 1 and 2 above. Under BSL’s apparent claim construction, Chen renders this claim obvious.

680. For example, Chen teaches that “[t]here are two LSP roles: primary or secondary (backup). The GMPLS signaling protocol carries a flag that indicates . . . the resource allocated for a backup LSP may be used by an LSP that has lower priority *until the primary LSP fails and the traffic is switched over to the backup.*” App. 7 (Chen) at pg. 21, last paragraph.

681. Thus, I agree with the Examiner that Chen discloses the additional elements of this claim.

682. In sum, it is my opinion that Chen renders claim 3 of the ’652 patent obvious in view of Voit and Blanchet under the Patent Owner’s proposed constructions of “priority” and “existing traffic.”

4. *Claim 4: A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.*

683. Claim 4 recites: “A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.”

684. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL’s apparent claim construction, Chen renders this claim obvious.

685. For example, Chen teaches that the backup path is dynamically chosen from a plurality of connections. App. 7 (Chen) at pg. 22.

686. Thus, I agree with the Examiner that Chen discloses the additional elements of this claim.

687. In sum, it is my opinion that Chen renders claim 4 of the ’652 patent obvious in view of Voit and Blanchet under the Patent Owner’s proposed constructions of “priority” and “existing traffic.”

5. *Claim 5: A method as recited in claim 1, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

688. Claim 5 recites: “A method as recited in claim 1, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.”

689. I incorporate by reference the portions of this declaration pertaining to Claim 1 above. Under BSL’s apparent claim construction, Chen renders this claim obvious.

690. For example, Chen discloses that “[d]uring LSP signaling in GMPLS, label distribution protocols may carry the link protection type. App. 7 (Chen) at pg. 41 ¶ 4.

691. Thus, I agree with the Examiner that Chen discloses the additional elements of this claim.

692. In sum, it is my opinion that Chen renders claim 5 of the '652 patent obvious in view of Voit and Blanchet under the Patent Owner's proposed constructions of "priority" and "existing traffic."

8. *Claim 8: A method as recited in claim 5, wherein the protection scheme indicates at least one of the following:*

693. Claim 8 recites: "A method as recited in claim 5, wherein the protection scheme indicates at least one of the following."

694. I incorporate by reference the portions of this declaration pertaining to Claims 1 and 5 above. Under BSL's apparent claim construction, Chen renders this claim obvious.

a. *a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

695. Claim 8 further recites: "a 1+1 protection scheme, wherein the same traffic is sent over two. . . ." Under BSL's apparent claim construction, Chen renders this element obvious.

696. For example, Chen discloses "dedicated 1+1" protection. App. 7 (Chen) at pg. 17.

b. *a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

697. Claim 8 further recites: "a 1:1 protection scheme, wherein one standby . . . is used to protect another . . . ." Under BSL's apparent claim construction, Chen renders this element obvious.

698. For example, Chen discloses "dedicated 1:1" protection. App. 7 (Chen) at pg. 17.

*c. a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or*

699. Claim 8 further recites: “a 1:1 protection scheme, wherein one standby . . . is used to protect another . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

700. For example, Chen discloses “1:N protection.” App. 7 (Chen) at pg. 54.

*d. an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.*

701. Claim 8 further recites: “an M:N protection scheme, wherein M standby . . . are used to protect N other . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

702. For example, Chen discloses “M:N protection.” App. 7 (Chen) at pg. 53.

703. Thus, I agree with the Examiner that Chen discloses the additional elements of this claim.

704. In sum, it is my opinion that Chen renders claim 8 of the ’652 patent obvious in view of Voit and Blanchet under the Patent Owner’s proposed constructions of “priority” and “existing traffic.”

**Claims 9-11, 13-15 and 17**

705. I note that the limitations of independent claim 9 and claim 14 are nearly identical to claim 1, except for the fact that claim 9 is a system claim and claim 14 is a computer program claim. The dependent claims also mirror the claims that depend on claim 1. For example, claim 10 is analogous to claim 2, claims 11 and 15 are analogous to claim 5, and claims 13 and 17 are analogous to claim 8. As such, my analysis below largely incorporates by reference my analysis with respect to claims 1-5 and 8. The claims correlate as follows:

<u>Method</u>	<u>System</u>	<u>Computer Program</u>
Claim 1	Claim 9	Claim 14
Claim 2	Claim 10	
Claim 3		
Claim 4		
Claim 5	Claim 11	Claim 15
Claim 8	Claim 13	Claim 17

9. *Claim 9: A system for providing protection to network traffic, comprising a processor configured to:*

706. Claim 9 recites: “a system for providing protection to network traffic, comprising a processor.” Under BSL’s apparent claim construction, Chen renders this claim obvious.

707. I incorporate by reference my comments from claim 1 above.

708. In addition, Chen discloses that “The original MPLS architecture [1] assumes that a Label Switching Router (LSR) has a forwarding plane which can (a) recognize packet (or cell) boundaries, and (b) process packet (or cell) headers. One skilled in the art would recognize that a “processor” must exist in order for an LSR to “process” a packet. App. 7 (Chen) at pg. 8.

709. Moreover, Chen discloses that the disclosed protection mechanisms are implemented on “routers” or “nodes.” A “router” or “node” on a network inherently includes a “processor.”

a. *send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property*

*associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;*

710. Claim 9 further recites: “send a . . . protection configuration parameter for configuring a standby . . . between a source node and a destination node, the . . . protection configuration parameter indicating a protection property associated with the standby . . ., the protection property including a priority for the standby . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

711. I incorporate by reference my comments from claim element 1(a) and 1(b) above.

*b. receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

712. Claim 9 further recites: “receive a . . . configuration acknowledgement indicating whether the . . . protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, Chen renders this element obvious.

713. I incorporate by reference my comments from claim element 1(c) above.

*c. accept the Pseudowire protection configuration parameter by the destination node;*

714. Claim 9 further recites: “accept the . . . protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, Chen renders this element obvious.

715. I incorporate by reference my comments from claim element 1(d) above.

*d. use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

716. Claim 9 further recites: “use the standby . . . that is configured based at least in part on the . . . protection configuration parameter.” Under BSL’s apparent claim construction, Chen renders this element obvious.

717. I incorporate by reference my comments from claim element 1(e) above.

*e. determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

718. Claim 9 further recites: “determine whether to preempt existing traffic on the standby . . . , wherein the determination is based, at least in part, on the priority for the standby . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

719. I incorporate by reference my comments from claim element 1(f) above.

720. In sum, it is my opinion that Chen renders claim 9 of the ’652 patent obvious in view of Voit and Blanchet.

*10. Claim 10: A system as recited in claim 9, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.*

721. Claim 10 recites: “A system as recited in claim 9, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.” Under BSL’s apparent claim construction, Chen renders this claim obvious.

722. I incorporate by reference my comments from claim 2 above.

723. In sum, it is my opinion that Chen renders claim 10 of the '652 patent obvious in view of Voit and Blanchet.

*11. Claim 11: A system as recited in claim 9, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

724. Claim 11 recites: "A system as recited in claim 9, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme." Under BSL's apparent claim construction, Chen renders this claim obvious.

725. I incorporate by reference my comments from claims 1, 5 and 9 above.

726. In sum, it is my opinion that Chen renders claim 11 of the '652 patent obvious in view of Voit and Blanchet.

*13. Claim 13: A system as recited in claim 11, wherein the protection scheme indicates at least one of the following:*

727. Under BSL's apparent claim construction, Chen renders this claim obvious.

728. I incorporate by reference my comments from claims 1, 5, 8, and 9 above.

*a. a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

729. Claim 11 further recites: "a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires." Under BSL's apparent claim construction, Chen renders this element obvious.



*b. a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

730. Claim 11 further recites: “a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire.” Under BSL’s apparent claim construction, Chen renders this element obvious.

731. I incorporate by reference my comments from claim 8 above.

*c. a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or*

732. Claim 11 further recites: “a 1:1 protection scheme, wherein one standby . . . is used to protect another . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

733. For example, Chen discloses “1:N protection.” App. 7 (Chen) at pg. 54.

734. I incorporate by reference my comments from claim 8 above.

*d. an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.*

735. Chen discloses “an M:N protection scheme, wherein M standby . . . are used to protect N other. . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

736. I incorporate by reference my comments from claim 8 above.

737. In sum, it is my opinion that Chen renders claim 13 of the ’652 patent obvious in view of Voit and Blanchet.

14. *Claim 14: A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for:*

738. Claim 14 recites: “A computer program product for configuring a [virtual path] between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions.” Under BSL’s apparent claim construction, Chen renders this claim obvious.

739. I incorporate by reference my comments from claim 1 above.

740. In addition, Chen teaches that, during configuration using RSVP protocol, “[e]ach router along the path *creates a software record (software state)* for the particular flow, which keeps the flow classifier, QoS requirements, next hops, previous hops and other related information.” App. 7 (Chen) at pg. 37.

741. As another example, Chen teaches that, when using the CR-LDP protocol to signal a path, “R5, as the ending node of the LSP, *programs* the label forwarding table, reserves the resource if needed, and responds with a CR-LDP Label Mapping message, which carries a Label TLV” and that “R1, as the head node of the LSP, does not need to allocate label any more, but simply receives the label and *programs* the label forwarding table.” App. 7 (Chen) at pgs. 49-50.

742. These disclosures in Chen suggest that the routers contain a “computer program” for configuring the virtual paths and that the programs are stored and run on each node.

a. *sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire*

*protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;*

743. Claim 14 further recites: “sending a . . . protection configuration parameter for configuring a standby . . . between a source node and a destination node, the . . . protection configuration parameter indicating a protection property associated with the standby . . ., the protection property including a priority for the standby . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

744. I incorporate by reference my comments from claim element 1(a) and 1(b) above.

*b. receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;*

745. Chen teaches “receiving a . . . configuration acknowledgement indicating whether the . . . protection configuration parameter has been accepted by the destination node.” Under BSL’s apparent claim construction, Chen renders this element obvious.

746. I incorporate by reference my comments from claim element 1(c) above.

*c. accept the Pseudowire protection configuration parameter by the destination node;*

747. Claim 14 further recites: “accepting the . . . protection configuration parameter by the destination node.” Under BSL’s apparent claim construction, Chen renders this element obvious.

748. I incorporate by reference my comments from claim element 1(d) above.

*d. using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and*

749. Claim 14 further recites: “using the standby . . . that is configured based at least in part on the . . . protection configuration parameter.” Under BSL’s apparent claim construction, Chen renders this element obvious.

750. I incorporate by reference my comments from claim element 1(e) above.

*e. determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.*

751. Claim 14 further recites: “determining whether to preempt existing traffic on the standby . . . , wherein the determination is based, at least in part, on the priority for the standby . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

752. I incorporate by reference my comments from claim element 1(f) above.

753. In sum, it is my opinion that Chen renders claim 14 of the ’652 patent obvious in view of Voit and Blanchet.

*15. Claim 15: A computer program product as recited in claim 14, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.*

754. Claim 15 recites: “A computer program product as recited in claim 14, wherein the protection property further includes at least one of a domain type, a

protection type or a protection scheme.” Under BSL’s apparent claim construction, Chen renders this claim obvious.

755. I incorporate by reference my comments from claims 1, 5 and 15 above.

756. In sum, it is my opinion that Chen renders claim 15 of the ’652 patent obvious in view of Voit and Blanchet.

17. *Claim 17: A computer product as recited in claim 15, wherein the protection scheme indicates at least one of the following:*

757. Claim 17 recites: “A computer product as recited in claim 15, wherein the protection scheme indicates at least one of the following.” Under BSL’s apparent claim construction, Chen renders this claim obvious.

758. I incorporate by reference my comments from claims 1, 5, 8, and 15 above.

a. *a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;*

759. Claim 17 further recites: “a 1+1 protection scheme, wherein the same traffic is sent over two . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

760. I incorporate by reference my comments from claim 8 above.

b. *a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;*

761. Claim 17 further recites: “a 1:1 protection scheme, wherein one standby . . . is used to protect another . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

762. I incorporate by reference my comments from claim 8 above.

*c. a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or*

763. Claim 17 further recites: “a 1:1 protection scheme, wherein one standby . . . is used to protect another . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

764. For example, Chen discloses “1:N protection.” App. 7 (Chen) at pg. 54.

765. I incorporate by reference my comments from claim 8 above.

*d. an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.*

766. Claim 17 further recites: “an M:N protection scheme, wherein M standby . . . are used to protect N other . . . .” Under BSL’s apparent claim construction, Chen renders this element obvious.

767. I incorporate by reference my comments from claim 8 above.

768. In sum, it is my opinion that Chen renders claim 17 of the ’652 patent obvious in view of Voit and Blanchet.

769. In sum, it is my opinion that Chen in view of Voit and Blanchet renders each of the Challenged Claims obvious under 35 U.S.C. § 103.

## VIII. SUMMARY OF MY OPINIONS

770. As discussed in further detail above, the alleged invention of the '652 patent was not novel. Instead, the concepts of Pseudowire signaling, protection configuration parameters, assigning “priorities,” and preempting traffic based on priorities were conventional and well-known.

771. Indeed, as shown by my analysis above, the concept of protecting the data on a Pseudowire using a “standby” or “backup” path was also well-known and a widely used technique in the industry and it was disclosed in detail in at least RFC 3386, Halabi, Chen, Owens, and Voit.

772. As further shown by my analysis above, at least Hofmeister, Halabi, RFC 3209 and Chen disclose at length the concepts of “sending a Pseudowire protection configuration parameter” between a source node and a destination node that contained a “protection property” was a standardized method of configuring a Pseudowire using the traditional signaling protocols of LDP and RSVP-TE. Similarly, the use of a “Pseudowire configuration acknowledgement” to indicate whether the requested protection parameters were accepted or rejected by the destination node was also a standard part of these protocols.

773. As further shown by my analysis above, it was common to assign a “priority” to working and protection Pseudowires by including a field for “priority” in the configuration parameter request message that is sent between a source node and destination node. Indeed, the concept of signaling priorities is disclosed in detail in at least Hofmeister, RFC 3386, Halabi, and RFC 3209.

774. As also shown by my analysis above, the concept of “preempting” existing traffic on a Pseudowire based on a “priority” that has been assigned to the Pseudowire to deal with bandwidth overloads and/or failures in the network was also conventional technology. This is demonstrated by the disclosures in at least Hofmeister, RFC 3386, Halabi, Chen, and Owens.

775. Based on my analysis above, it is my opinion that claims 1-5, 8-11, 13-15, and 17 of the '652 patent are anticipated by many prior art references, including **U.S. Patent Pub. No. 2004/0156313** to Hofmeister et al., **Request for Comments 3386** and/or **“Metro Ethernet”** by Sam Halabi.

776. If certain aspects recited in claims 1-5, 8-11, 13-15, and 17 of the '652 patent are not deemed to be disclosed or inherent over these references, then claims 1-5, 8-11, 13-15, and 17 of the '652 patent are certainly obvious in view of some combination of these references and or in combination with **U.S. Patent No. 7,804,767 B1** to Owens et al., **Request for Comments 3209**, **“The LSP Protection/Restoration Mechanism in GMPLS”** by Ziyang Chen, **U.S. Patent No. 7,305,481 B2** to Blanchet et al. and/or **U.S. Patent Pub. No. 2006/0047851 A1** to Voit et al.

777. The bases for my opinions are set forth in detail above.



## IX. CONCLUSION

For the reasons stated herein, it is my opinion that the Challenged Claims of the '652 patent are anticipated or obvious. This declaration is based on my present assessment of materials and information currently available to me. My investigation and assessment may continue, which may include reviewing documents and other information that may yet to be made available to me. Accordingly, I expressly reserve the right to continue my study in connection with this case and to expand or modify my opinions and conclusions as my study continues.

I declare under penalty of perjury under the laws of the United States that the foregoing is true and correct.

Respectfully submitted,

Dated: 2/10/2014

By: Tal Lavian  
TAL LAVIAN, Ph.D.

# Tal Lavian, Ph.D.



<http://innovations-IP.com>  
<http://cs.berkeley.edu/~tlavian>  
[tlavian@innovations-IP.com](mailto:tlavian@innovations-IP.com)

1640 Mariani Dr.  
Sunnyvale, CA 94087  
(408)-209-9112

---

## Research and Consulting: Network Communications, Telecommunications, and Internet Software

- Scientist, educator, and technologist with over 25 years-of experience
- Co-author of over 25 scientific publications, journal articles, and peer-reviewed papers
- Named inventor on over 80 issued and filed patents
- Industry Fellow and Lecturer at UC Berkeley Engineering – Center for Entrepreneurship and Technology (CET)

---

## EDUCATION

- **Ph.D.**, Computer Science specializing in networking and communications, UC Berkeley
- **M.Sc.**, Electrical Engineering, Tel Aviv University
- **B.Sc.**, Mathematics and Computer Science, Tel Aviv University

## TECHNOLOGIES

Network communications, telecommunications, and Internet software technologies:

- **Communication networks**: TCP/IP suite, TCP, UDP, IP, VoIP, Ethernet, Data Link, ARP, ICMP, network protocols, network software applications
- **Mobile Wireless**: Wireless LAN, cellular systems, mobile devices, smartphone technologies
- **Routing/switching**: LAN, WAN, VPN, routing protocols, RIP, BGP, MPLS, DNS, QoS, NAP, switching, packet switching, network infrastructure, network communication architectures
- **Internet Software**: Internet software applications, Internet protocols, distributed computing, Web applications, FTP, HTTP, Java, C, C++, Client Server, file transfer, multicast, streaming media

## PROFESSIONAL SUMMARY

- Selected as Principal Investigator for three US Department of Defense (DARPA) projects
- Led research project on networking computation for the US Air Force Research Lab (AFRL)
- Led and developed the first network resource scheduling service for grid computing
- Led wireless research project for an undisclosed US federal agency
- Managed and engineered the first demonstrated transatlantic dynamic allocation of 10Gbs Lambdas as a grid service
- Spearheaded and planned the first demonstrated wire-speed active network on commercial hardware
- Created and chaired Nortel Networks' EDN Patent Committee
- IEEE Senior Member

## PROFESSIONAL EXPERIENCE

**Innovations-IP**, Sunnyvale, CA

2006-Present

### **Principal Scientist**

- Consults in the areas of network communications, telecommunications, Internet software technologies, and smartphone mobile wireless devices
- CTO at VisuMenu, a very small stealth stage company developing visual IVR technologies for smartphones and wireless mobile devices in the area of network communications (since 2010)
- Provides architecture and system consultation for software projects relating to mobile wireless devices, Internet web applications, and computer networks
- Expert witness in network communications patent infringement suits

**University of California Berkeley**, Berkeley, CA

2000-Present

### **Berkeley Industry Fellow, Lecturer, Visiting Scientist, Ph.D. Candidate, Nortel's Scientist Liaison**

*Some positions and projects were concurrent, others sequential*

- Serves as Industry Fellow and Lecturer at the Center for Entrepreneurship and Technology (CET)
- Studies the areas of network services, telecommunication systems and software, communications infrastructure, and data centers
- Developed long-term technology for the enterprise market, integrating communication and computing technologies
- Conducted research projects in data centers (RAD Labs), telecommunication infrastructure (SAHARA), and wireless systems (ICEBERG)
- Acted as scientific liaison between Nortel Research Lab and UC Berkeley, providing tangible value in advanced technologies
- Earned Ph.D. in Computer Science, specializing in communications and networking

**Nortel Networks**, Santa Clara, CA

1996 - 2007

### **Principal Scientist, Principal Architect, Principal Engineer, Senior Software Engineer**

- Held scientific and research roles at Nortel Labs, Bay Architecture Labs, and CTO Office

### **-Principal Investigator for US Department of Defense (DARPA) Projects**

- Conceived, proposed, and completed three research projects: Active Networks, DWDM-RAM, and a networking computation project for Air Force Research Lab (AFRL)
- Led a wireless research project for an undisclosed US federal agency

### **-Academic and Industrial Researcher**

- Analyzed new technologies with the objective of reducing risks associated with R&D investment
- Spearheaded research collaboration with leading universities and professors at UC Berkeley, Northwestern University, University of Amsterdam, and University of Technology Sydney
- Evaluated competitive products relative to Nortel's products and technology

- Proactively identified prospective business ideas, leading to new networking products
- Predicted technological trends well in advance through researching the technological horizon and academic sphere
- Developed software for switches, routers and network communications devices
- Developed systems and architectures for switches, routers, and network management
- Researched and developed the following projects:
  - Data-Center Communications: network and server orchestration 2006-2007
  - DRAC: SOA-facilitated L1/L2/L3 network dynamic controller 2003-2007
  - Omega: classified wireless project for undisclosed US Federal Agency 2006
  - Open Platform: project for the US Air Force Research Laboratory (AFRL) 2005
  - Network Resource Orchestration for Web Services Workflows 2004-2005
  - Proxy Study between Web/Grids Services and Network Services 2004
  - Streaming Content Replication: real-time A/V media multicast at edge 2003-2004
  - DWDM-RAM: US DARPA-funded program on agile optical transport 2003-2004
  - Packet Capturing and Forwarding Service on IP and Ethernet traffic 2002-2003
  - CO2: content-aware agile networking 2001-2003
  - Active Networks: US DARPA-funded research program 1999-2002
  - ORE: programmable network service platform 1998-2002
  - JVM Platform: Java on network devices 1998-2001
  - Web-Based Device Management: network device management 1996-1997

**-Technology Innovator and Patent Leader**

- Created and chaired Nortel Networks' EDN Patent Committee
- Facilitated continuous stream of innovative ideas and their conversion into intellectual property rights
- Developed intellectual property assets through invention and analysis of existing technology portfolios

**Aptel Communications**, Netanya, Israel 1994-1995

**Software Engineer, Team Leader**

*Start-up company focused on mobile wireless CDMA spread spectrum PCN/PCS*

- Developed mobile wireless device using an unlicensed band, Direct Sequence Spread Spectrum (DSSS)
- Designed and managed a personal communication network (PCN) and personal communication system (PCS), the precursors of short text messages (SMS)
- Responsible for the design and development of network software products
- Developed software network communications mainly in C/C++
- Brought two-way paging product from concept to development

**Scitex Ltd.**, Herzeliya, Israel 1990-1993

**Software Engineer, Team Leader**

*Software and hardware company acquired by Hewlett Packard (HP)*

- Developed system and network communications mainly in C/C++
- Invented Parallel SIMD Architecture
- Participated in the Technology Innovation group

*Start-up company*

**Software Engineer**

- Developed real-time software and algorithms mainly in C/C++ and Pascal

**PROFESSIONAL ASSOCIATIONS**

- IEEE Senior Member
- IEEE CNSV co-chair Intellectual Property SIG (2013)
- President Next Step Toastmasters (the only advanced TM club in the Silicon Valley) (2013)
- Technical Co-Chair, IEEE Hot Interconnects 2005 at Stanford University
- Member, IEEE Communications Society (COMMSOC)
- Member, IEEE Computer Society
- Member, IEEE Systems, Man, and Cybernetics Society
- Member, IEEE-USA Intellectual Property Committee
- Member, ACM, ACM Special Interest Group on Data Communication (SIGCOM)
- Member, ACM Special Interest Group on Hypertext, Hypermedia and Web (SIGWEB)
- Member, IEEE Consultants' Network (CNSV)
- Global Member, Internet Society (ISOC)
- President Java Users Group – Silicon Valley Mountain View, CA, 1999-2000
- Toastmasters International

**ADVISORY BOARDS**

- Quixey – search engine for wireless mobile apps
- Mytopia – mobile social games
- iLeverage – Israeli Innovations

**PROFESSIONAL AWARDS**

- Top Talent Award – Nortel
- Top Inventors Award – Nortel EDN
- Certified IEEE-WCET - Wireless Communications Engineering Technologies
- Toastmasters International - Competent Communicator (twice)
- Toastmasters International - Advanced Communicator Bronze

## Patents and Publications

*(not an exhaustive list)*

### Patents Issued

- **US 8,537,989** Device and method for providing enhanced telephony
- **US 8,078,708** Grid proxy architecture for network resources
- **US 7,944,827** Content-aware dynamic network resource allocation
- **US 7,860,999** Distributed computation in network devices
- **US 7,734,748** Method and apparatus for intelligent management of a network element
- **US 7,710,871** Dynamic assignment of traffic classes to a priority queue in a packet forwarding device
- **US 7,580,349** Content-aware dynamic network resource allocation
- **US 7,433,941** Method and apparatus for accessing network information on a network device
- **US 7,359,993** Method and apparatus for interfacing external resources with a network element
- **US 7,313,608** Method and apparatus for using documents written in a markup language to access and configure network elements
- **US 7,260,621** Object-oriented network management interface
- **US 7,237,012** Method and apparatus for classifying Java remote method invocation transport traffic
- **US 7,127,526** Method and apparatus for dynamically loading and managing software services on a network device
- **US 7,047,536** Method and apparatus for classifying remote procedure call transport traffic
- **US 7,039,724** Programmable command-line interface API for managing operation of a network device
- **US 6,976,054** Method and system for accessing low-level resources in a network device
- **US 6,970,943** Routing architecture including a compute plane configured for high-speed processing of packets to provide application layer support
- **US 6,950,932** Security association mediator for Java-enabled devices
- **US 6,850,989** Method and apparatus for automatically configuring a network switch
- **US 6,845,397** Interface method and system for accessing inner layers of a network protocol
- **US 6,842,781** Download and processing of a network management application on a network device
- **US 6,772,205** Executing applications on a target network device using a proxy network device
- **US 6,564,325** Method of and apparatus for providing multi-level security access to system
- **US 6,175,868** Method and apparatus for automatically configuring a network switch
- **US 6,170,015** Network apparatus with Java co-processor
- **US 8,406,388** Systems and methods for visual presentation and selection of IVR menu
- **US 8,155,280** Systems and methods for visual presentation and selection of IVR menu
- **US 8,054,952** Systems and methods for visual presentation and selection of IVR menu
- **US 8,000,454** Systems and methods for visual presentation and selection of IVR menu
- **US 8,223,931** Systems and methods for visual presentation and selection of IVR menu
- **US 8,160,215** Systems and methods for visual presentation and selection of IVR menu
- **EP 1,905,211** Technique for authenticating network users
- **EP 1,142,213** Dynamic assignment of traffic classes to a priority queue in a packet forwarding device
- **EP 1,671,460** Method and apparatus for scheduling resources on a switched underlay network
- **CA 2,358,525** Dynamic assignment of traffic classes to a priority queue in a packet forwarding device
- **US 8,161,139** Method and apparatus for intelligent management of a network element
- **US 8,146,090** Time-value curves to provide dynamic QoS for time sensitive file transfer
- **US 8,341,257** Grid proxy architecture for network resource
- **US 8,345,835** Systems and methods for visual presentation and selection of IVR menu

## Patent Applications Published and Pending

- **US 20130080898** SYSTEMS AND METHODS FOR ELECTRONIC COMMUNICATIONS
- **US 20130022191** SYSTEMS AND METHODS FOR VISUAL PRESENTATION AND SELECTION OF IVR MENU
- **US 20130022183** SYSTEMS AND METHODS FOR VISUAL PRESENTATION AND SELECTION OF IVR MENU
- **US 20130022181** SYSTEMS AND METHODS FOR VISUAL PRESENTATION AND SELECTION OF IVR MENU
- **US 20120180059** TIME-VALUE CURVES TO PROVIDE DYNAMIC QoS FOR TIME SENSITIVE FILE TRANSFERS
- **US 20120063574** SYSTEMS AND METHODS FOR VISUAL PRESENTATION AND SELECTION OF IVR MENU
- **US 20110225330** PORTABLE UNIVERSAL COMMUNICATION DEVICE
- **US 20100220616** OPTIMIZING NETWORK CONNECTIONS
- **US 20100217854** Method and Apparatus for Intelligent Management of a Network Element
- **US 20100146492** TRANSLATION OF PROGRAMMING CODE
- **US 20100146112** EFFICIENT COMMUNICATION TECHNIQUES
- **US 20100146111** EFFICIENT COMMUNICATION IN A NETWORK
- **US 20090313613** Methods and Apparatus for Automatic Translation of a Computer Program Language Code
- **US 20090313004** Platform-Independent Application Development Framework
- **US 20090279562** Content-aware dynamic network resource allocation
- **US 20080040630** Time-Value Curves to Provide Dynamic QoS for Time Sensitive File Transfers
- **US 20070169171** Technique for authenticating network users
- **US 20060123481** Method and apparatus for network immunization
- **US 20060075042** Extensible resource messaging between user applications and network elements in a communication network
- **US 20050083960** Method and apparatus for transporting parcels of data using network elements with network element storage
- **US 20050076339** Method and apparatus for automated negotiation for resources on a switched underlay network
- **US 20050076336** Method and apparatus for scheduling resources on a switched underlay network
- **US 20050076173** Method and apparatus for preconditioning data to be transferred on a switched underlay network
- **US 20050076099** Method and apparatus for live streaming media replication in a communication network
- **US 20050074529** Method and apparatus for transporting visualization information on a switched underlay network
- **US 20040076161** Dynamic assignment of traffic classes to a priority queue in a packet forwarding device
- **US 20020021701** Dynamic assignment of traffic classes to a priority queue in a packet forwarding device

## Publications

(not an exhaustive list)

- “Communications Architecture in Support of Grid Computing”, Tal Lavian, Scholar's Press 2013 ISBN 978-3-639-51098-0.
- “Applications Drive Secure Lightpath Creation across Heterogeneous Domains, Feature Topic Optical Control Planes for Grid Networks: Opportunities, Challenges and the Vision.” Gommans L.; Van Oudenaarde B.; Dijkstra F.; De Laat C.; Lavian T.; Monga I.; Taal A.; Travostino F.; Wan A.; *IEEE Communications Magazine*, vol. 44, no. 3, March 2006, pp. 100-106.
- *Lambda Data Grid: Communications Architecture in Support of Grid Computing*. Tal I. Lavian, Randy H. Katz; Doctoral Thesis, University of California at Berkeley. January 2006.
- “Information Switching Networks.” Hoang D.B.; T. Lavian; *The 4th Workshop on the Internet, Telecommunications and Signal Processing, WITSP 2005*, December 19-21, 2005, Sunshine Coast, Australia.
- “Impact of Grid Computing on Network Operators and HW Vendors.” Allcock B.; Arnaud B.; Lavian T.; Papadopoulos P.B.; Hasan M.Z.; Kaplow W.; *IEEE Hot Interconnects at Stanford University 2005*, pp.89-90.
- *DWDM-RAM: A Data Intensive Grid Service Architecture Enabled by Dynamic Optical Networks*. Lavian T.; Mambretti J.; Cutrell D.; Cohen H.J.; Merrill S.; Durairaj R.; Daspit P.; Monga I.; Naiksatam S.; Figueira S.; Gutierrez D.; Hoang D.B., Travostino F.; *CCGRID 2004*, pp. 762-764.
- *DWDM-RAM: An Architecture for Data Intensive Service Enabled by Next Generation Dynamic Optical Networks*. Hoang D.B.; Cohen H.; Cutrell D.; Figueira S.; Lavian T.; Mambretti J.; Monga I.; Naiksatam S.; Travostino F.; *Proceedings IEEE Globecom 2004, Workshop on High-Performance Global Grid Networks*, Houston, 29 Nov. to 3 Dec. 2004, pp.400-409.
- *Implementation of a Quality of Service Feedback Control Loop on Programmable Routers*. Nguyen C.; Hoang D.B.; Zhao, I.L.; Lavian, T.; *Proceedings, 12th IEEE International Conference on Networks 2004. (ICON 2004) Singapore, Volume 2, 16-19 Nov. 2004*, pp.578-582.
- *A Platform for Large-Scale Grid Data Service on Dynamic High-Performance Networks*. Lavian T.; Hoang D.B.; Mambretti J.; Figueira S.; Naiksatam S.; Kaushil N.; Monga I.; Durairaj R.; Cutrell D.; Merrill S.; Cohen H.; Daspit P.; Travostino F.; *GridNets 2004, San Jose, CA., October 2004*.
- *DWDM-RAM: Enabling Grid Services with Dynamic Optical Networks*. Figueira S.; Naiksatam S.; Cohen H.; Cutrell D.; Daspit, P.; Gutierrez D.; Hoang D. B.; Lavian T.; Mambretti J.; Merrill S.; Travostino F.; *Proceedings, 4th IEEE/ACM International Symposium on Cluster Computing and the Grid, Chicago, USA, April 2004*, pp. 707-714.
- *DWDM-RAM: Enabling Grid Services with Dynamic Optical Networks*. Figueira S.; Naiksatam S.; Cohen H.; Cutrell D.; Gutierrez D.; Hoang D.B.; Lavian T.; Mambretti J.; Merrill S.; Travostino F.; *4th IEEE/ACM International Symposium on Cluster Computing and the Grid, Chicago, USA, April 2004*.
- *An Extensible, Programmable, Commercial-Grade Platform for Internet Service Architecture*. Lavian T.; Hoang D.B.; Travostino F.; Wang P.Y.; Subramanian S.; Monga I.; *IEEE Transactions on Systems, Man, and*



Cybernetics on Technologies Promoting Computational Intelligence, Openness and Programmability in Networks and Internet Services Volume 34, Issue 1, Feb. 2004, pp.58-68.

- *DWDM-RAM: An Architecture for Data Intensive Service Enabled by Next Generation Dynamic Optical Networks*. Lavian T.; Cutrell D.; Mambretti J.; Weinberger J.; Gutierrez D.; Naiksatam S.; Figueira S.; Hoang D. B.; Supercomputing Conference, SC2003 Igniting Innovation, Phoenix, November 2003.
- *Edge Device Multi-Unicasting for Video Streaming*. Lavian T.; Wang P.; Durairaj R.; Hoang D.; Travostino F.; Telecommunications, 2003. ICT 2003. 10th International Conference on Telecommunications, Tahiti, Volume 2, 23 Feb.-1 March, 2003 pp. 1441-1447.
- *The SAHARA Model for Service Composition Across Multiple Providers*. Raman B.; Agarwal S.; Chen Y.; Caesar M.; Cui W.; Lai K.; Lavian T.; Machiraju S.; Mao Z. M.; Porter G.; Roscoe T.; Subramanian L.; Suzuki T.; Zhuang S.; Joseph A. D.; Katz Y.H.; Stoica I.; Proceedings of the First International Conference on Pervasive Computing. ACM Pervasive 2002, pp. 1-14.
- *Enabling Active Flow Manipulation in Silicon-Based Network Forwarding Engines*. Lavian T.; Wang P.; Travostino F.; Subramanian S.; Durairaj R.; Hoang D.B.; Sethaput V.; Culler D.; Proceeding of the Active Networks Conference and Exposition, 2002.(DANCE) 29-30 May 2002, pp. 65-76.
- *Practical Active Network Services within Content-Aware Gateways*. Subramanian S.; Wang P.; Durairaj R.; Rasimas J.; Travostino F.; Lavian T.; Hoang D.B.; Proceeding of the DARPA Active Networks Conference and Exposition, 2002.(DANCE) 29-30 May 2002, pp. 344-354.
- *Active Networking on a Programmable Network Platform*. Wang P.Y.; Lavian T.; Duncan R.; Jaeger R.; Fourth IEEE Conference on Open Architectures and Network Programming (OPENARCH), Anchorage, April 2002.
- *Intelligent Network Services through Active Flow Manipulation*. Lavian T.; Wang P.; Travostino F.; Subramanian S.; Hoang D.B.; Sethaput V.; IEEE Intelligent Networks 2001 Workshop (IN2001), Boston, May 2001.
- *Intelligent Network Services through Active Flow Manipulation*. Lavian T.; Wang P.; Travostino F.; Subramanian S.; Hoang D.B.; Sethaput V.; Intelligent Network Workshop, 2001 IEEE 6-9 May 2001, pp.73 - 82.
- *Enabling Active Flow Manipulation in Silicon-based Network Forwarding Engine*. Lavian, T.; Wang, P.; Travostino, F.; Subramanian S.; Hoang D.B.; Sethaput V.; Culler D.; Journal of Communications and Networks, March 2001, pp.78-87.
- *Active Networking on a Programmable Networking Platform*. Lavian T.; Wang P.Y.; IEEE Open Architectures and Network Programming, 2001, pp. 95-103.
- *Enabling Active Networks Services on a Gigabit Routing Switch*. Wang P.; Jaeger R.; Duncan R.; Lavian T.; Travostino F.; 2nd Workshop on Active Middleware Services, 2000.

- *Dynamic Classification in Silicon-Based Forwarding Engine Environments*. Jaeger R.; Duncan R.; Travostino F.; Lavian T.; Hollingsworth J.; Selected Papers. 10th IEEE Workshop on Metropolitan Area and Local Networks, 1999. 21-24 Nov. 1999, pp.103-109.
- *Open Programmable Architecture for Java-Enabled Network Devices*. Lavian, T.; Jaeger, R. F.; Hollingsworth, J. K.; IEEE Hot Interconnects Stanford University, August 1999, pp. 265-277.
- *Open Java SNMP MIB API*. Rob Duncan, Tal Lavian, Roy Lee, Jason Zhou, Bay Architecture Lab Technical Report TR98-038, December 1998.
- *Java-Based Open Service Interface Architecture*. Lavian T.; Lau S.; BAL TR98-010 Bay Architecture Lab Technical Report, March 1998.
- *Parallel SIMD Architecture for Color Image Processing*. Lavian T. Tel – Aviv University, Tel – Aviv, Israel, November 1995.

## Presentations and Talks

*(not an exhaustive list)*

- Lambda Data Grid: An Agile Optical Platform for Grid Computing and Data-intensive Applications.
- Web Services and OGSA
- WINER Workflow Integrated Network Resource Orchestration.
- Technology & Society.
- Abundant Bandwidth and how it affects us?
- Active Content Networking(ACN).
- DWDM-RAM:Enabling Grid Services with Dynamic Optical Networks .
- Application-engaged Dynamic Orchestration of Optical Network Resources .
- A Platform for Data Intensive Services Enabled by Next Generation Dynamic Optical Networks .
- Optical Networks.
- Grid Optical Network Service Architecture for Data Intensive Applications.
- Optical Networking & DWDM.
- OptiCal Inc.
- OptiCal & LUMOS Networks.
- Optical Networking Services.
- Business Models for Dynamically Provisioned Optical Networks.
- Business Model Concepts for Dynamically Provisioned Optical Networks.
- Optical Networks Infrastructure.
- Research Challenges in agile optical networks.
- Services and Applications' infrastructure for agile optical networks.
- Impact on Society.
- TeraGrid Communication and Computation.
- Unified Device Management via Java-enabled Network Devices.
- Active Network Node in Silicon-Based L3 Gigabit Routing Switch.
- Active Nets Technology Transfer through High-Performance Network Devices.
- Programmable Network Node: Applications.
- Open Innovation via Java-enabled Network Devices.
- Practical Considerations for Deploying a Java Active Networking Platform.
- Open Java-Based Intelligent Agent Architecture for Adaptive Networking Devices.
- Java SNMP Oplet.
- Open Distributed Networking Intelligence: A New Java Paradigm.
- Open Programmability.
- Active Networking On A Programmable Networking Platform.
- Open Networking through Programmability.
- Open Programmable Architecture for Java-enabled Network Devices.
- Integrating Active Networking and Commercial-Grade Routing Platforms.
- Programmable Network Devices.
- To be smart or not to be?





US007940652B1

(12) **United States Patent**  
**Pan**

(10) **Patent No.:** **US 7,940,652 B1**  
(45) **Date of Patent:** **May 10, 2011**

(54) **PSEUDOWIRE PROTECTION USING A  
STANDBY PSEUDOWIRE**

(75) Inventor: **Ping Pan**, San Jose, CA (US)

(73) Assignee: **Brixham Solutions Ltd.**, Tortola (VG)

(\* ) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 861 days.

(21) Appl. No.: **11/354,569**

(22) Filed: **Feb. 14, 2006**

**Related U.S. Application Data**

(60) Provisional application No. 60/653,065, filed on Feb. 14, 2005.

(51) **Int. Cl.**  
**H04J 3/14** (2006.01)

(52) **U.S. Cl.** ..... **370/228; 370/216; 370/225; 709/220**

(58) **Field of Classification Search** ..... **370/216, 370/225, 228; 709/220**

See application file for complete search history.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

5,920,705 A	7/1999	Lyon et al.	
6,167,051 A	12/2000	Nagami et al.	
6,347,088 B1	2/2002	Katou et al.	
6,430,184 B1	8/2002	Robins et al.	
6,546,427 B1	4/2003	Ehrlich	
6,574,477 B1 *	6/2003	Rathunde	455/453
6,621,793 B2	9/2003	Widegren et al.	
6,665,273 B1	12/2003	Goguen et al.	
6,680,943 B1	1/2004	Gibson et al.	
6,751,684 B2	6/2004	Owen et al.	
6,813,271 B1	11/2004	Cable	
6,845,389 B1	1/2005	Sen et al.	
6,985,488 B2	1/2006	Pan et al.	
7,050,396 B1	5/2006	Cohen et al.	
7,200,104 B2 *	4/2007	Saleh et al.	370/216

7,436,782 B2	10/2008	Ngo et al.	
7,697,528 B2	4/2010	Parry	
2001/0021175 A1	9/2001	Haverinen	
2001/0023453 A1	9/2001	Sundqvist	
2002/0112072 A1	8/2002	Jain	
2002/0141393 A1	10/2002	Eriksson	
2002/0146026 A1	10/2002	Unitt et al.	
2003/0002482 A1	1/2003	Kubler et al.	
2003/0039237 A1	2/2003	Forslow	
2003/0117950 A1 *	6/2003	Huang	370/220
2004/0105459 A1	6/2004	Mannam	
2004/0114595 A1	6/2004	Doukai	
2004/0133692 A1 *	7/2004	Blanchet et al.	709/230
2004/0156313 A1	8/2004	Hofmeister	
2004/0174865 A1	9/2004	O'Neill	
2004/0252717 A1	12/2004	Solomon et al.	

(Continued)

**OTHER PUBLICATIONS**

Ziying Chen: "The LSP Protection/Restoration Mechanism in GMPLS" Internet Citatio (Online) Oct. 2002 (Oct. 1, 2002), XP002239552 Retrieved from the ineternet URL: <http://www.site.uottawa.ca/~bochmann/dsrg/PublicDocuments/Master-theses/Chen,%20Ziying%20-%20-%202002.pdf>\*

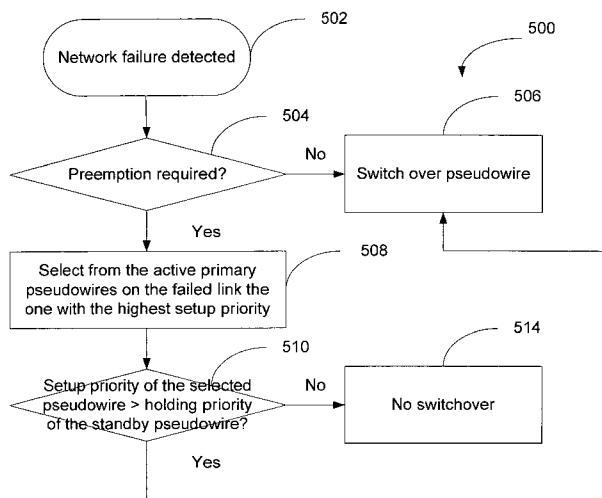
(Continued)

*Primary Examiner* — William Trost, IV  
*Assistant Examiner* — Siming Liu

(57) **ABSTRACT**

Providing protection to network traffic includes sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node, and in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire, wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter.

**17 Claims, 7 Drawing Sheets**



U.S. PATENT DOCUMENTS

2005/0018605	A1	1/2005	Foote	
2005/0044262	A1	2/2005	Luo	
2005/0220148	A1	10/2005	DelRegno	
2005/0237927	A1	10/2005	Kano et al.	
2006/0002423	A1	1/2006	Rembert	
2006/0018252	A1 *	1/2006	Sridhar et al.	370/216
2006/0046658	A1 *	3/2006	Cruz et al.	455/67.11
2006/0047851	A1 *	3/2006	Voit et al.	709/239
2006/0090008	A1	4/2006	Guichard	
2006/0146832	A1	7/2006	Rampal	
2006/0233167	A1	10/2006	McAllister	
2007/0053366	A1	3/2007	Booth, III	
2007/0127479	A1	6/2007	Sinicrope et al.	
2007/0206607	A1	9/2007	Chapman	
2008/0031129	A1	2/2008	Arseneault	

OTHER PUBLICATIONS

Braden, R. et al., "Integrated Services in the Internet Architecture: an overview," Network Working Group, Jun. 1994.

Bryant, S. et al., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture," Network Working Group, Mar. 2005.

Blake, S. et al., "An Architecture for Differentiated Services," Network Working Group, Dec. 1998.

Shah, Himanshu et al., Internet Draft, ARP Mediation for IP Interworking of Layer 2 VPN, L2VPN Working Group, Jul. 2007.

Martini, Luca et al., Internet Draft, Segmented Pseudo Wire, Network Working Group, Jul. 2007.

Pan, P. et al., Internet Draft, Pseudo Wire Protection, Jul. 2006.

Rosen, Eric C. et al., Internet Draft, PWE3 Congestion Control Framework, Network Working Group, Mar. 2004.

Rosen, E. et al., BGO-MPLS IP Virtual Private Networks (VPN), Network Working Group, Feb. 2006.

Pan, Ping, Internet Draft, Dry-Martini: Supporting Pseudo-wires in Sub-IP Access Networks, Network Working Group, Jul. 2005.

Mcperson et al., Pseudowire Emulation Edge to Edge (PWE3) Jun. 13, 2007, <http://www.ietf.org/html.charters/pwe3-carter.html>.

Afferton, Thomas S. et al., Ethernet Transport over Wide Area Networks, Packet-Aware Transport for Metro Networks, IEEE Communications Magazine, pp. 120-127, Mar. 2004.

Martini, L. et al., Pseudowire Setup and Maintenance using the Label Distribution Protocol (LDP), Network Working Group, Apr. 2006.

Anderson, L. et al., LDP Specification, Network Working Group, Jan. 2001.

Martini, Luca et al., Encapsulation Methods for Transport of Ethernet over MPLS Networks, Network Working Group, Apr. 2006.

Martini, Luca et al., Encapsulation Methods for Transport of Frame Relay Over MPLS Networks, Network Working Group, Feb. 2006.

Metz, Chris et al., Pseudowire Attachment Identifiers for Aggregation and VPN Autodiscovery, PWE3 Working Group, Feb. 25, 2006.

Martini, Luca et al., Dynamic Placement of Multi Segment Pseudo Wires, PWE3 Working Group, Jun. 2006.

Martini, Luca et al., "Pseudowire Setup and Maintenance using LDP", Network Working Group, Mar. 2005.

Vasseur, et al., Path Computation Element (pce), May 9, 2007, <http://www.ietf.org/html.charters/pce.charter.html>.

Theimer, T. et al., "Requirements for OAM Functionality in MPLS", Oct. 1999, Watersprings.

Harry Newton, "Newton's Telecom Dictionary", 23rd Updated and Expanded Edition, p. 825, p. 239, Flatiron Publishing, New York, Mar. 2007.

\* cited by examiner

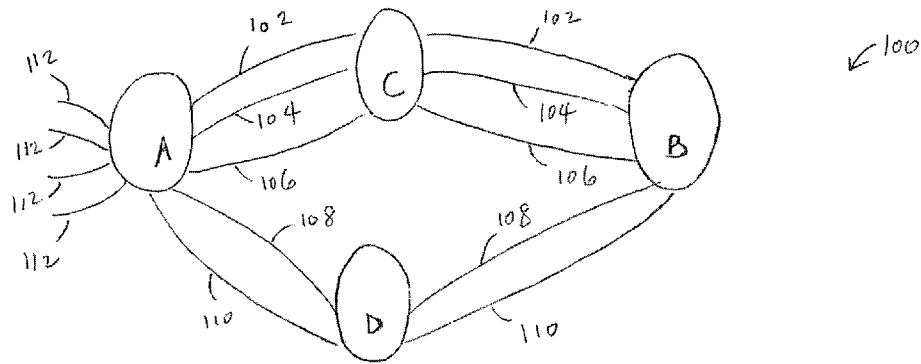


FIG. 1A

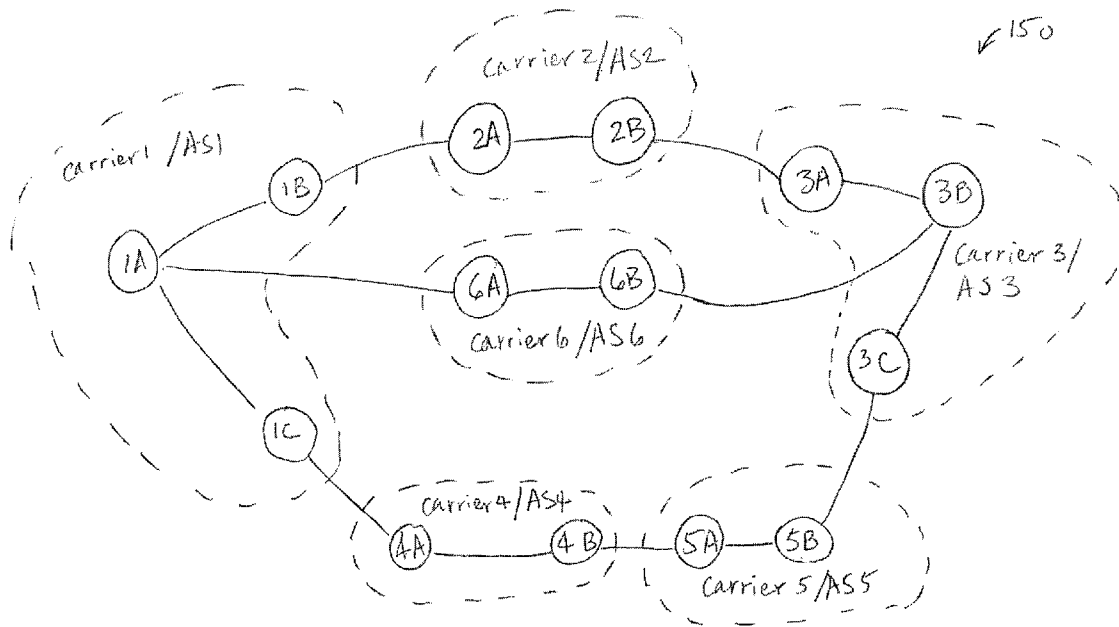


FIG. 1B

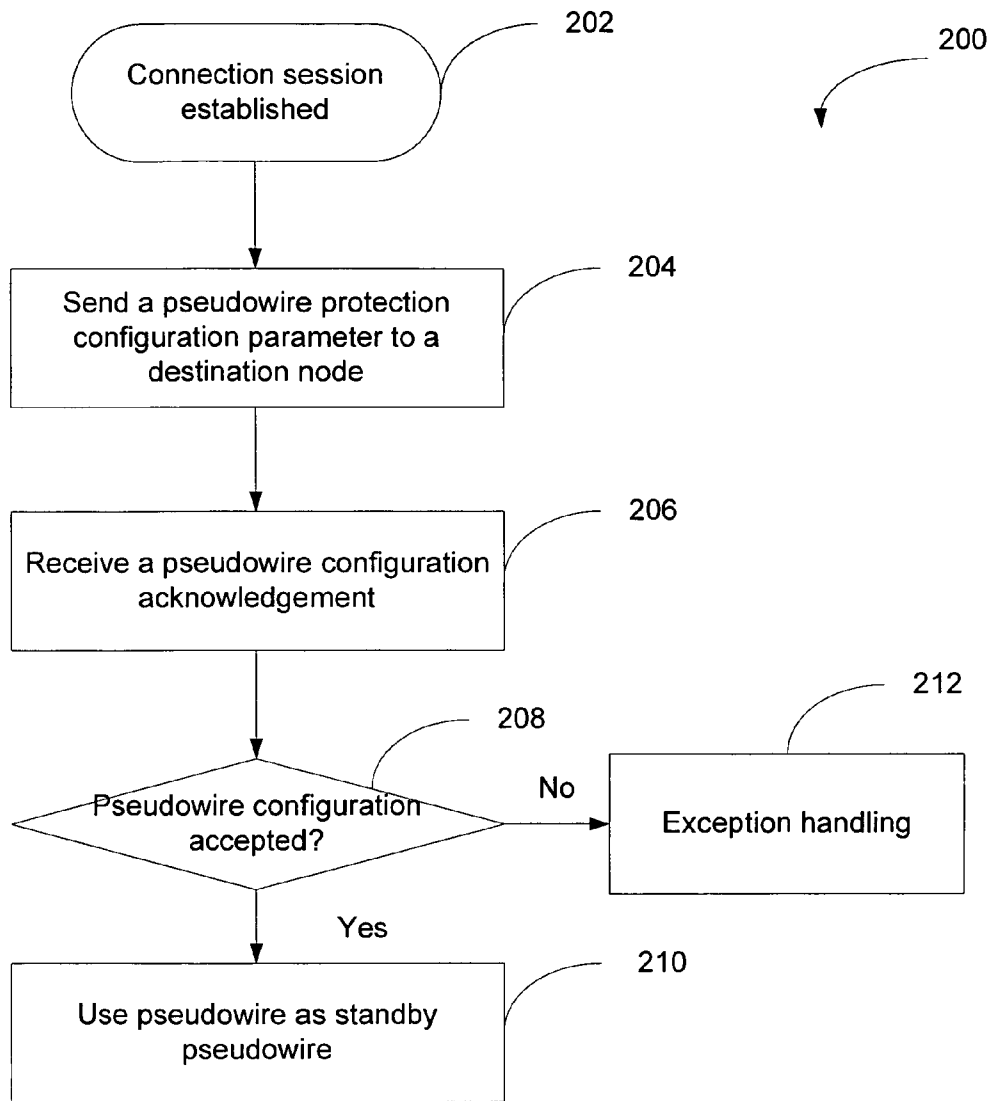


FIG. 2



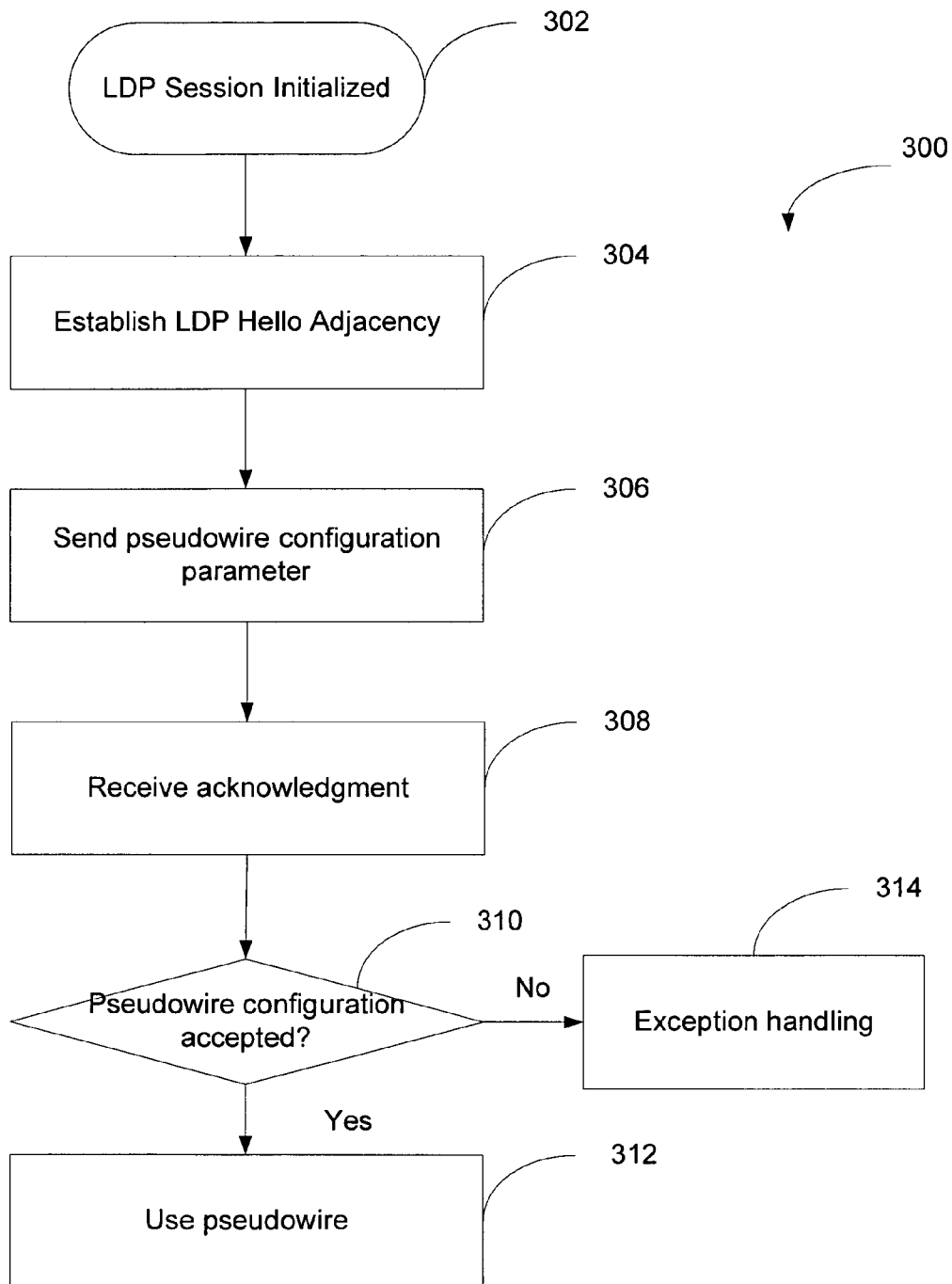


FIG. 3A

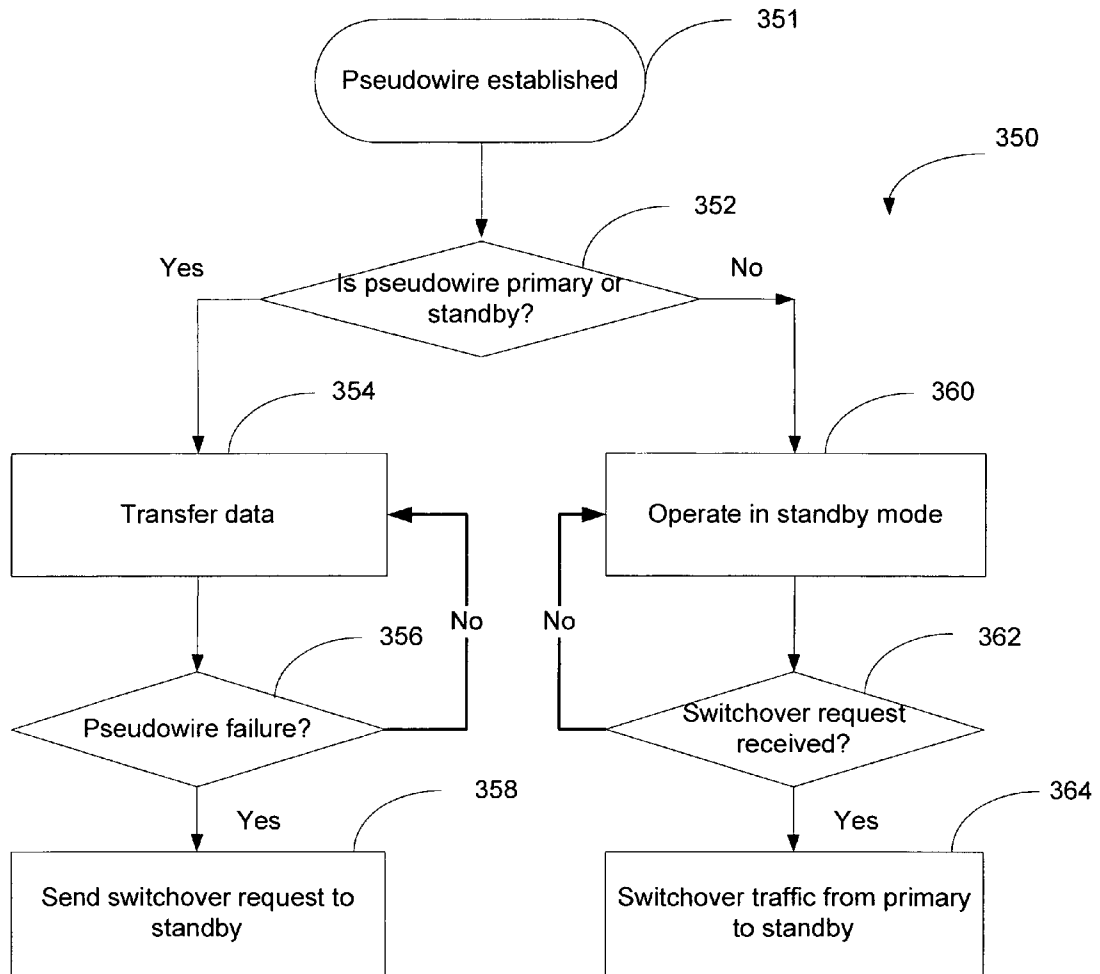


FIG. 3B

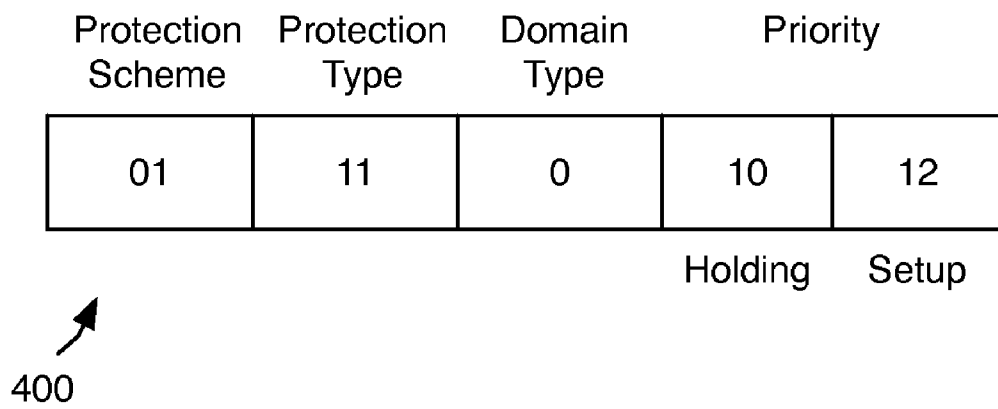


FIG. 4

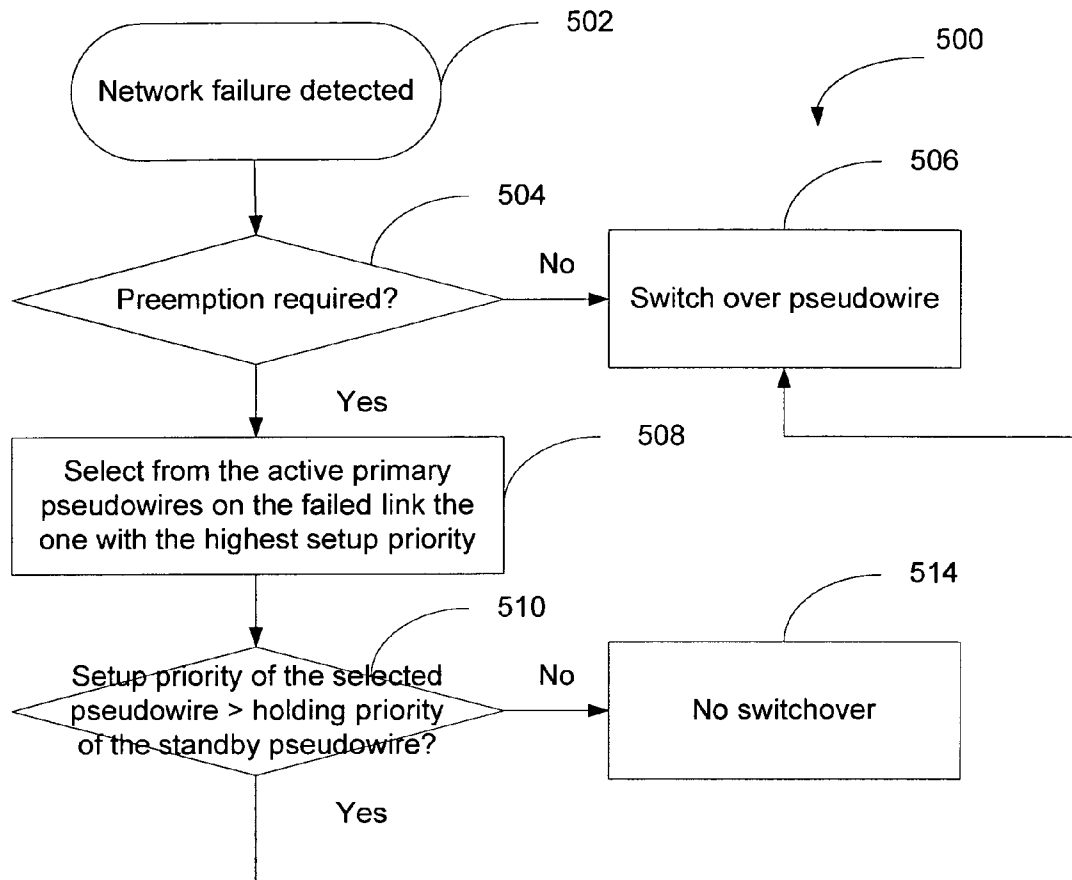


FIG. 5

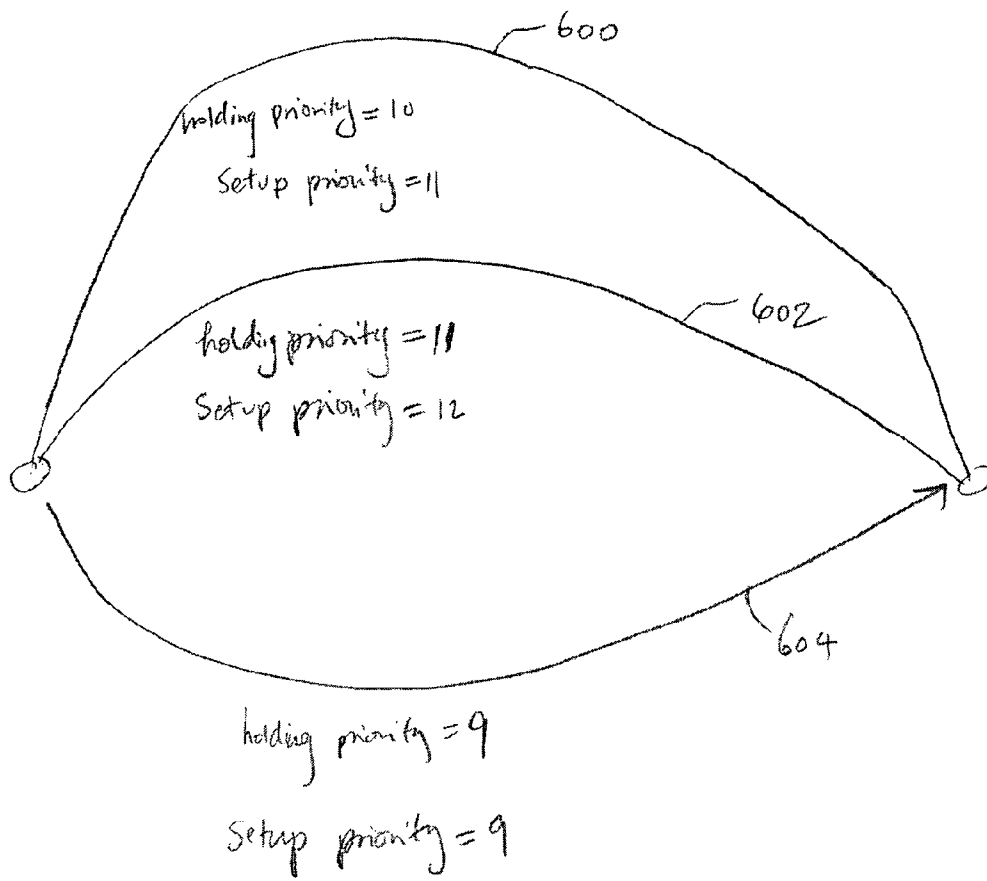


FIG. 6

1

## PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE

### CROSS REFERENCE TO OTHER APPLICATIONS

This application claims priority to U.S. Provisional Patent Application No. 60/653,065 entitled PSEUDO WIRE PROTECTION filed Feb. 14, 2005 which is incorporated herein by reference for all purposes.

### BACKGROUND OF THE INVENTION

In recent years, many networking and telecommunications carriers have deployed Pseudowires to carry Layer-2 (also known as the data link layer of the Open Systems Interconnection (OSI) Reference Model) traffic. A Pseudowire (PW) refers to an emulation of a native service over a network. Examples of the native service include Asynchronous Transfer Mode (ATM), Frame Relay, Ethernet, Time Division Multiplexing (TDM), Synchronous Optical Network (SONET), Synchronous Digital Hierarchy (SDH), etc. Examples of the network include Multiprotocol Label Switching (MPLS), Internet Protocol (IP), etc. More recently, a number of carriers have extended the use of Pseudowires beyond packet encapsulation, and offered Pseudowires as a type of network service. Consequently, data traffic protection and redundancy in environments that use Pseudowire have become critical.

At the edge of a network, a network edge device such as an edge router may receive multiple Layer-2 flows (also referred to as Attachment Circuits (ACs)). In a typical network supporting Pseudowires, each AC is mapped to a Pseudowire. Ingress packets received mapped to a specific Pseudowire are labeled with an identifier associated with this Pseudowire, and are switched via the Pseudowire. A physical link may support one or more Pseudowires. Ideally, the data flow in a Pseudowire should be protected. In other words, if an active Pseudowire fails, the data flow should be redirected to an alternative Pseudowire to avoid data loss.

Pseudowires can operate over many physical media types. However, existing Pseudowire systems typically provide no protection or very limited protection. For example, there is usually no data protection for Pseudowires on different physical media types, since most network protection schemes, such as APS for SONET, Link Aggregation for Ethernet, do not apply over multiple physical media types.

Some MPLS devices implement schemes such as MPLS Fast Reroute to provide limited data protection. These existing schemes, however, often do not provide adequate protection. Take the following scenario as an example: between two provider edges (PEs), a first tunnel comprising multiple Pseudowires is protected by a second tunnel. Due to network topology constraints, the two tunnels may have different bandwidth. This is a possible scenario in an MPLS Fast Reroute operation. In this example, the second tunnel may have lower bandwidth than that of the first one. If the first tunnel should fail, the amount of data that needs to be redirected through the second tunnel may exceed the capacity of the second tunnel. Furthermore, existing protocols typically do not provide a way of determining which data gets priority. Thus, certain mission critical data may be dropped while other less critical data may pass through.

It would be desirable to have a way to provide better Pseudowire protection and to have more control during switchover. It would also be desirable if the protection

2

scheme could be implemented without significant changes to existing protocols and devices.

### BRIEF DESCRIPTION OF THE DRAWINGS

Various embodiments of the invention are disclosed in the following detailed description and the accompanying drawings.

FIGS. 1A and 1B are block diagrams illustrating an embodiment of a single-hop Pseudowire system and an embodiment of a multi-hop Pseudowire system, respectively.

FIG. 2 is a flowchart illustrating an embodiment of a process of providing data protection using Pseudowires.

FIG. 3A is a flowchart illustrating another embodiment of a process of providing data protection using Pseudowires.

FIG. 3B is a flowchart illustrating how the Pseudowire is used, according to some embodiments.

FIG. 4 is a data structure diagram illustrating an embodiment of a Pseudowire protection configuration parameter that specifies several protection-related properties of the Pseudowire.

FIG. 5 is a flowchart illustrating an example process of using the priorities during switchover.

FIG. 6 is a diagram illustrating an example in which pre-emption takes place during a switchover operation.

### DETAILED DESCRIPTION

The invention can be implemented in numerous ways, including as a process, an apparatus, a system, a composition of matter, a computer readable medium such as a computer readable storage medium or a computer network wherein program instructions are sent over optical or electronic communication links. In this specification, these implementations, or any other form that the invention may take, may be referred to as techniques. A component such as a processor or a memory described as being configured to perform a task includes both a general component that is temporarily configured to perform the task at a given time or a specific component that is manufactured to perform the task. In general, the order of the steps of disclosed processes may be altered within the scope of the invention.

A detailed description of one or more embodiments of the invention is provided below along with accompanying figures that illustrate the principles of the invention. The invention is described in connection with such embodiments, but the invention is not limited to any embodiment. The scope of the invention is limited only by the claims and the invention encompasses numerous alternatives, modifications and equivalents. Numerous specific details are set forth in the following description in order to provide a thorough understanding of the invention. These details are provided for the purpose of example and the invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the invention is not unnecessarily obscured.

Providing protection to network traffic using one or more Pseudowires is disclosed. In some embodiments, a Pseudowire protection configuration parameter is sent to a destination node. A Pseudowire configuration acknowledgment from the destination node is received. If a Pseudowire is allowed to be established according to the Pseudowire configuration acknowledgment, it is established based at least in part on the Pseudowire protection configuration parameter. In embodiments where the Pseudowire is established as a standby

Pseudowire configured to protect one or more primary Pseudowires, in the event that a primary Pseudowire fails to transfer network traffic for reasons such as network congestion, equipment failure, etc., network traffic that is originally designated to be transferred on the primary Pseudowire(s) is switched from the primary Pseudowire(s) to the standby Pseudowire.

The protection technique is applicable to both single-hop and multi-hop systems. FIGS. 1A and 1B are block diagrams illustrating an embodiment of a single-hop Pseudowire system and an embodiment of a multi-hop Pseudowire system, respectively. Configuring and switching the Pseudowire will be discussed in more detail below.

In the example shown in FIG. 1A, system 100 is a single-hop system where the nodes in the system all belong to the same carrier network. Within each carrier network, all network nodes and facility are under a common administrative control. A service provider company may own multiple carrier networks in different regions. As used herein, a node refers to a networked device. In this case, the nodes in the system are provider edges (PEs) A, B, C, and D, which all belong to the same carrier network. Ingress data received by attachment circuits 112 of PE A designated for PE B may be sent via a label switched path (LSP) through PEs A, C, and B, or an LSP through PEs A, D, and B. The first LSP comprises Pseudowires 102, 104 and 106, and the second LSP comprises Pseudowires 108 and 110. In this example, the Pseudowire connections between PEs are established using the Label Distribution Protocol (LDP). The connections are based on LDP sessions. Each LDP session is to connect two local or remote nodes. There may be multiple paths interconnecting any two nodes in the network. Thus, for each LDP session, there may be multiple LDP Hello Adjacencies, one LDP Hello Adjacency per path. For purposes of example, throughout this specification, LDP is used as the communication protocol between nodes. Other appropriate protocols may also be used.

In the example shown in FIG. 1B, system 150 is a multi-hop system since it includes multiple carrier networks. Carrier networks 1-6 form autonomous systems 1-6, respectively. Each autonomous system includes one or more networks that are controlled by a carrier. For purposes of illustration, three Pseudowires are shown in this example to transfer data between PE 1A and PE 3B: a first Pseudowire comprising a path via autonomous systems 1, 2, and 3, a second Pseudowire comprising a path via autonomous systems 1, 6, and 3, and a third Pseudowire comprising a path via autonomous systems 1, 4, 5, and 3. Other Pseudowire formations are possible. At the source node PE 1A, data packets to be sent via a particular Pseudowire are labeled with an identifier associated with the Pseudowire, forwarded on to the next provider edge on one Pseudowire segment, and forwarded again if necessary until the packets reach the destination node 3B.

FIG. 2 is a flowchart illustrating an embodiment of a process of providing data protection using Pseudowires. Process 200 may be implemented on a source node such as A or 1A of systems 100 and 150, or on an independent management agent that communicates with the source node. For purposes of illustration, the process is shown as implemented on a source node in the following example. The process initializes when a connection session is established between the source node and the destination node (202). A Pseudowire protection configuration parameter for configuring a Pseudowire based on the connection session is sent (204). The Pseudowire protection configuration parameter includes one or more fields that specify certain protection properties associated with the Pseudowire. It may be sent to the destination node or a man-

agement agent that communicates with the destination node. Details of the configuration parameter will be discussed further below.

Once the destination node (or its associated management agent) receives the Pseudowire protection configuration parameter, it determines whether it will accept the Pseudowire protection configuration and allow a standby Pseudowire to be established. Depending on implementation, the destination node determines whether to accept the protection configuration based on factors such as traffic condition, number of existing Pseudowires, priority information, etc. The destination node may reject the protection request for a number of reasons. For example, the destination node does not support Pseudowire protection mechanism as described here. If a standby Pseudowire may be established, the destination node accepts it and configures the Pseudowire based at least in part on the configuration parameters. In some embodiments, the destination node adds the Pseudowire to a table of Pseudowires. A corresponding Pseudowire configuration acknowledgment is generated, indicating whether the destination node has accepted the Pseudowire configuration. The Pseudowire configuration acknowledgment is sent to the source node. In some embodiments, as a part of the LDP process, a MPLS label for the data packets traversing through the standby Pseudowire is assigned.

At the source node, once the Pseudowire configuration acknowledgment is received (206), it is examined to determine whether the Pseudowire configuration has been accepted (208). If, according to the Pseudowire configuration acknowledgment, the Pseudowire configuration has been accepted by the destination, a standby Pseudowire is established based at least in part on the Pseudowire protection configuration parameter and may be used as such (210). If, however, the Pseudowire configuration has not been accepted, the process performs appropriate exception handling, such as re-sending the Pseudowire protection configuration parameter (212).

FIG. 3A is a flowchart illustrating another embodiment of a process of providing data protection using Pseudowires. Process 300 may be implemented on a PE, on an independent management agent, or the like. For purposes of illustration, in the following example, the process is initiated and carried out on a PE source node.

Process 300 begins with the initialization of an LDP session (302). According to the negotiation scheme based on LDP, the source node exchanges messages with the destination node and establishes an LDP Hello Adjacency (304). A Pseudowire setup request that includes a Pseudowire protection configuration parameter is sent to the destination node (or its associated management agent), requesting that a standby Pseudowire be established over the LDP Hello Adjacency (306). In some embodiments, multiple LDP Hello Adjacencies are available for Pseudowire setup, thus multiple setup requests are sent, and the destination node processes the requests and maps Pseudowires to appropriate LDP Hello Adjacencies. In some embodiments, the source node dynamically determines which LDP Hello Adjacency among the available connections is to be configured as a standby Pseudowire, and directs its setup request accordingly. The dynamic determination may be based on, among other things, bandwidth availability on the adjacency path.

In some embodiments, the request is sent as a LDP Label Mapping Message. The configuration parameter is used to configure various properties of the Pseudowire, including protection type, protection scheme, priority, etc. Further details of the configuration parameters are discussed below. In some embodiments, multiple LDP Hello Adjacencies are

5

established and the source node sends multiple Pseudowire setup requests to configure Pseudowires over these LDP Hello Adjacencies.

In this example, upon receiving a Pseudowire setup request, the destination node maps the request to the appropriate LDP Hello Adjacency. If the mapping is successful, the Pseudowire is established. Sometimes, however, the mapping and consequently the Pseudowire setup may fail for reasons such as network congestion, resource limitation, equipment failure, etc. The destination node sends a Pseudowire configuration acknowledgment to the source node. In this example, the Pseudowire configuration acknowledgment is an LDP acknowledgement indicating whether a particular Pseudowire has been successfully established. Once the source node receives the acknowledgement (308), it determines whether the configuration has been accepted by the destination (310). If the configuration has been accepted, a standby Pseudowire is successfully established based at least in part on the Pseudowire protection configuration parameter, and the source and destination nodes can start using the standby Pseudowire to protect other Pseudowires (312). If, however, the acknowledgment indicates that the configuration has not been accepted and a Pseudowire has not been successfully established, appropriate exception handling measures such as resending the Pseudowire protection configuration parameter are taken (314).

Process 300 is applicable to both single-hop and multi-hop systems. In a single-hop system, the source node and the destination node correspond to a source PE and a destination PE on the network and the process is used to configure a standby Pseudowire between the PEs. In a multi-hop system, the process may be repeated by the PEs on various carrier networks to establish Pseudowire segments. For example, in system 150 of FIG. 1B, PE 1A can use process 300 to establish a Pseudowire segment with PE 6A, and PE 6A can use the same process to establish a Pseudowire segment with PE 6B, which can use the same process to establish a Pseudowire segment with PE 3B.

FIG. 3B is a flowchart illustrating how the Pseudowire is used, according to some embodiments. Process 350 may be implemented on the source node, the destination node, or both. In this example, the designation of the Pseudowire is first determined (352). The designation may be configured by a system administrator, in a Pseudowire configuration process, or any other appropriate means. If the Pseudowire is designated as a primary Pseudowire, it is configured to carry network traffic (354). In the event that a primary Pseudowire fails (356), the nodes associated with the Pseudowire will attempt to switch the traffic over to the standby Pseudowire by sending a switchover request to the Pseudowire (358). As will be shown in more detail below, in some embodiments, whether the traffic on the primary Pseudowire can preempt the traffic on the standby Pseudowire and be switched over depends on priority configuration of the Pseudowires.

If it is designated as a standby Pseudowire, it enters into standby mode to provide protection to one or more primary Pseudowires (360). In some embodiments, the standby Pseudowire carries network traffic during normal operation. It is ready to take over traffic from the primary Pseudowire if necessary. If a switchover request is received from a primary Pseudowire (362), traffic on the primary Pseudowire is switched over to the standby Pseudowire. In some embodiments, the switchover only occurs if the priority comparison of the primary and standby Pseudowires indicates the switchover is allowed.

Optionally, during the operation, if a Pseudowire is no longer needed, the source node can send a withdraw request

6

over the Pseudowire and the destination node disassociates the Pseudowire with the LDP Hello Adjacency to break the Pseudowire connection.

FIG. 4 is a data structure diagram illustrating an embodiment of a Pseudowire protection configuration parameter that specifies several protection-related properties of the Pseudowire. In this example, Pseudowire protection configuration parameter 400 includes four fields: protection scheme, protection type, domain type, and priority. A field may have one or more subfields. For example, the priority field is shown to include a holding priority and a setup priority. One or more of the fields and/or subfields may be used in various embodiments. Other appropriate fields may also be implemented. In the example shown, the fields are numerical values that map to appropriate property values.

In some embodiments, one of the following Pseudowire protection schemes is used to set up the Pseudowires: 1+1, 1:1, 1:N or M:N. The protection scheme field is used to indicate which protection scheme is used in the system setup. A specific protection scheme corresponds to a field value. For example, 1+1 maps to 0, 1:1 maps to 1, and so on. In a system implementing a 1+1 protection scheme, the same traffic is sent over two parallel Pseudowires and the receiver selects one traffic stream at a time. In a system implementing a 1:1 protection scheme, one Pseudowire is used to protect another Pseudowire. Similarly, in a 1:N system (e.g. MPLS Facility Backup), one Pseudowire is used to protect N other Pseudowires, and in a M:N system M Pseudowires are used to protect N other Pseudowires.

The protection type field is used to configure the standby mode of the Pseudowire. In some embodiments, cold, warm, and hot standby modes are supported. Other appropriate standby modes may be implemented in other embodiments. In some embodiments, in cold standby mode configuration, once network failure on a Pseudowire carrying network traffic is detected, a standby Pseudowire is selected from the remaining functional Pseudowires, and traffic is redirected to the standby Pseudowire. In some embodiments with warm standby mode configuration, one or more standby Pseudowires are established before any network failure has occurred. These standby Pseudowires, however, are not maintained or used to transport data until a network failure is detected. Upon failure detection, the source or destination nodes will modify the data-plane and switch data traffic over to the standby Pseudowire(s). In some embodiments with hot standby mode configuration, one or more standby Pseudowires are pre-established and maintained at both control-plane and data-plane, so that once a network failure is detected, data traffic is directly switched over to the standby Pseudowire(s).

The domain type field indicates whether the Pseudowire is configured in a single-hop environment where all the nodes of the Pseudowire belong to the same carrier network, or a multi-hop environment where the Pseudowire includes nodes on several carrier networks. This is because the intermediate may process single-hop and multi-hop Pseudowire differently.

The priority field indicates the preference level of a Pseudowire in preempting other Pseudowires during switchover. In the event of a network failure, the edge nodes will preferentially provide protection according to the priority setting of the Pseudowires. In a situation where network resources (such as bandwidth) are limited, data sent on a higher priority Pseudowire is more likely to be protected than data sent on a lower priority Pseudowire. In some embodiments, the priority field includes two subfields: a holding priority and a setup priority. The holding priority indicates the relative priority of a currently active Pseudowire with respect



7

to other Pseudowires when the latter attempt to preempt the former's use of the data link. Stated another way, it determines how easily a currently active Pseudowire gives up its hold on a data link upon request. The setup priority indicates the relative priority of a Pseudowire during the setup process.

FIG. 5 is a flowchart illustrating an example process of using the priorities during switchover. Process 500 may be implemented on an edge node, an independent management agent, or the like. In this example, process 500 initiates when a network failure has been detected (502). It is determined whether preemption is required (504). Preemption is required when the failed link carries more Pseudowire traffic than the available bandwidth on the standby link. If preemption is not required, the Pseudowire(s) may directly switchover (506). If, however, preemption is required, the setup priorities of the Pseudowires on the failed link are compared and the Pseudowire with the highest setup priority is selected (508). The setup priority of the selected Pseudowire is compared to the holding priority of the standby Pseudowire (510). If the setup priority is greater than the holding priority, traffic on the selected Pseudowire is switched over to the standby Pseudowire (506). If, however, the setup priority is no greater than the holding priority, no switchover takes place and the standby Pseudowire continues to transfer its own data and the data on the failed Pseudowires is lost (514).

FIG. 6 is a diagram illustrating an example in which preemption takes place during a switchover operation. In this example, Pseudowires 600, 602 and 604 are active, primary Pseudowires carrying traffic. Pseudowire 604 is used as the standby Pseudowire. Pseudowire 600 has a holding priority and a setup priority of 10 and 11, respectively, Pseudowire 602 has priorities of 11 and 12, and Pseudowire 604 has priorities of 9 and 9. Thus, if the link on which Pseudowires 600 and 602 operate fails, the nodes will initiate switchover using Pseudowire 604. A comparison of the setup priority of Pseudowires 600 and 602 indicates that Pseudowire 602 has a higher setup priority, thus 602 is given preference in the switchover. The setup priority of Pseudowire 602 is compared with the holding priority of Pseudowire 604. Since 602's setup priority is greater than 604's holding priority, data on 602 preempts data on 604 and takes over the link.

Providing protection to network traffic using one or more Pseudowires has been disclosed. Pseudowire protection improves the reliability of Pseudowire services. Pseudowires are better controlled by appropriately configuring the properties of Pseudowires and without requiring significant changes to existing protocols and devices.

Although the foregoing embodiments have been described in some detail for purposes of clarity of understanding, the invention is not limited to the details provided. There are many alternative ways of implementing the invention. The disclosed embodiments are illustrative and not restrictive.

What is claimed is:

1. A method of providing protection to network traffic, comprising:

sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;

receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;

8

accepting the Pseudowire protection configuration parameter by the destination node;

using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and

determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

2. A method as recited in claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.

3. A method as recited in claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.

4. A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.

5. A method as recited in claim 1, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.

6. A method as recited in claim 1, wherein the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP).

7. A method as recited in claim 5, wherein the domain type indicates whether the standby Pseudowire is configured in a single-hop environment where the standby Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the standby Pseudowire includes a plurality of nodes coupled to several carrier networks.

8. A method as recited in claim 5, wherein the protection scheme indicates at least one of the following:

a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;

a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;

a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or

an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.

9. A system for providing protection to network traffic, comprising:

a processor configured to:

send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;

receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;

accept the Pseudowire protection configuration parameter by the destination node;

use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and

determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

9

10. A system as recited in claim 9, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.

11. A system as recited in claim 9, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.

12. A system as recited in claim 11, wherein the domain type indicates whether the standby Pseudowire is configured in a single-hop environment where the standby Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the standby Pseudowire includes a plurality of nodes coupled to several carrier networks.

13. A system as recited in claim 11, wherein the protection scheme indicates at least one of the following:

a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;

a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;

a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or

an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.

14. A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for:

sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;

10

receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;

accept the Pseudowire protection configuration parameter by the destination node;

using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and

determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

15. A computer program product as recited in claim 14, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.

16. A computer product as recited in claim 15, wherein the domain type indicates whether the standby Pseudowire is configured in a single-hop environment where the standby Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the standby Pseudowire includes a plurality of nodes coupled to several carrier networks.

17. A computer product as recited in claim 15, wherein the protection scheme indicates at least one of the following:

a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;

a 1:1 protection scheme, wherein one standby Pseudowire is used to protect another Pseudowire;

a 1:N protection scheme, wherein one standby Pseudowire is used to protect N other Pseudowires; or

an M:N protection scheme, wherein M standby Pseudowires are used to protect N other Pseudowires.

\* \* \* \* \*



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	05/10/2011	7940652	002.P045	6912

65638 7590 04/20/2011  
OMIKRON IP LAW GROUP  
16325 Boones Ferry Rd.  
SUITE 204  
LAKE OSWEGO, OR 97035

**ISSUE NOTIFICATION**

The projected patent number and issue date are specified above.

**Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)**  
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 861 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (<http://pair.uspto.gov>).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site <http://pair.uspto.gov> for additional applicants):

Ping Pan, San Jose, CA;

6-9. (Canceled).

10. (Original) A method as recited in claim 1, wherein the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP).

11. (Previously Presented) A system for providing protection to network traffic, comprising:

a processor configured to:

send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;

receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;

accept the Pseudowire protection configuration parameter by the destination node;

use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and

determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

12. (Original) A system as recited in Claim 11, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.

13. (Previously Presented) A system as recited in Claim 11, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.

14-~~17~~. (Canceled).  
16

Change(s) applied  
to document,  
/S.P.E./  
4/12/2011



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	002.P045	6912
65638	7590	03/14/2011	EXAMINER	
OMIKRON IP LAW GROUP 16325 Boones Ferry Rd. SUITE 204 LAKE OSWEGO, OR 97035			LIU, SIMING	
			ART UNIT	PAPER NUMBER
			2472	
			MAIL DATE	DELIVERY MODE
			03/14/2011	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Response to Rule 312 Communication</b>	<b>Application No.</b>	<b>Applicant(s)</b>
	11/354,569	PAN, PING
	<b>Examiner</b>	<b>Art Unit</b>
	SIMING LIU	2472

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**

1.  The amendment filed on 01 March 2011 under 37 CFR 1.312 has been considered, and has been:
- a)  entered.
  - b)  entered as directed to matters of form not affecting the scope of the invention.
  - c)  disapproved because the amendment was filed after the payment of the issue fee.  
Any amendment filed after the date the issue fee is paid must be accompanied by a petition under 37 CFR 1.313(c)(1) and the required fee to withdraw the application from issue.
  - d)  disapproved. See explanation below.
  - e)  entered in part. See explanation below.

/Hassan Kizou/  
Supervisory Patent Examiner, Art Unit 2472

/S. L./  
Examiner, Art Unit 2472

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No.: 11/354,569	Confirmation No.: 6912
Applicant: Ping Pan	Group Art Unit: 2472
Filing Date: February 14, 2006	Examiner: Liu, Siming
Docket No.: 002.P045	
Customer No.: 65638	<b>AMENDMENT AFTER NOTICE OF ALLOWANCE MAILED DECEMBER 2, 2010</b>
For: Pseudowire Protection Using a Standby Pseudowire	<b>SUBMITTED THROUGH EFS-WEB</b>

**AMENDMENT AFTER ALLOWANCE**

**PURSUANT TO 37 C.F.R. § 1.312**

Dear Examiner:

Please amend the application as indicated on the following pages.

Amendments to the Claims begin on page 2 of this paper.

Remarks begin at page 7 of this paper.

**AMENDMENTS TO THE CLAIMS:**

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Previously Presented) A method of providing protection to network traffic, comprising:
  - sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;
  - receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;
  - accepting the Pseudowire protection configuration parameter by the destination node;
  - using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and
  - determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.
2. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.
3. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.
4. (Original) A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.
5. (Previously Presented) A method as recited in claim 1, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.



6-9. (Canceled).

10. (Original) A method as recited in claim 1, wherein the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP).

11. (Previously Presented) A system for providing protection to network traffic, comprising:

a processor configured to:

send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;

receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;

accept the Pseudowire protection configuration parameter by the destination node;

use the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and

determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

12. (Original) A system as recited in Claim 11, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.

13. (Previously Presented) A system as recited in Claim 11, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.

14-17. (Canceled).

17. (Previously Presented) A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for:

    sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;

    receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node;

    accepting the Pseudowire protection configuration parameter by the destination node; using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and

    determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

18. (Previously Presented) A computer program product as recited in claim 17, wherein the protection property further includes at least one of a domain type, a protection type or a protection scheme.

19-21. (Canceled).

22. (Currently Amended) A method as recited in claim 5, wherein the domain type indicates whether the **standby** Pseudowire is configured in a single-hop environment where the **standby** Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the **standby** Pseudowire includes a plurality of nodes coupled to several carrier networks.

23. (Currently Amended) A method as recited in claim 5, wherein the protection scheme indicates at least one of the following:

    a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;

    a 1:1 protection scheme, wherein one **standby** Pseudowire is used to protect another Pseudowire;

a 1:N protection scheme, wherein one **standby** Pseudowire is used to protect N other Pseudowires; or

an M:N protection scheme, wherein M **standby** Pseudowires are used to protect N other Pseudowires.

24. (Currently Amended) A system as recited in claim 13, wherein the domain type indicates whether the **standby** Pseudowire is configured in a single-hop environment where the **standby** Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the **standby** Pseudowire includes a plurality of nodes coupled to several carrier networks.

25. (Currently Amended) A system as recited in claim 13, wherein the protection scheme indicates at least one of the following:

a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;

a 1:1 protection scheme, wherein one **standby** Pseudowire is used to protect another Pseudowire;

a 1:N protection scheme, wherein one **standby** Pseudowire is used to protect N other Pseudowires; or

an M:N protection scheme, wherein M **standby** Pseudowires are used to protect N other Pseudowires.

26. (Currently Amended) A computer product as recited in claim 18, wherein the domain type indicates whether the **standby** Pseudowire is configured in a single-hop environment where the **standby** Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the **standby** Pseudowire includes a plurality of nodes coupled to several carrier networks.

27. (Currently Amended) A computer product as recited in claim 18, wherein the protection scheme indicates at least one of the following:

a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;

a 1:1 protection scheme, wherein one **standby** Pseudowire is used to protect another Pseudowire;

a 1:N protection scheme, wherein one **standby** Pseudowire is used to protect N other Pseudowires; or

an M:N protection scheme, wherein M **standby** Pseudowires are used to protect N other Pseudowires.

**REMARKS**

Applicant has amended dependent claims 22-27 to provide proper antecedent basis support. No new matter has been added. *The Examiner is respectfully requested to contact the undersigned by telephone at (503) 551-9442 if the Examiner has any questions.*

Respectfully submitted,

Date: March 1, 2011

by: /Ted A. Crawford/Reg. No. 50,610/  
Ted A. Crawford  
Reg. No. 50,610

## Electronic Acknowledgement Receipt

<b>EFS ID:</b>	9564054
<b>Application Number:</b>	11354569
<b>International Application Number:</b>	
<b>Confirmation Number:</b>	6912
<b>Title of Invention:</b>	PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE
<b>First Named Inventor/Applicant Name:</b>	Ping Pan
<b>Customer Number:</b>	65638
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt
<b>Filer Authorized By:</b>	Ted A. Crawford
<b>Attorney Docket Number:</b>	002.P045
<b>Receipt Date:</b>	01-MAR-2011
<b>Filing Date:</b>	14-FEB-2006
<b>Time Stamp:</b>	18:15:41
<b>Application Type:</b>	Utility under 35 USC 111(a)

### Payment information:

Submitted with Payment	no
------------------------	----

### File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Amendment after Notice of Allowance (Rule 312)	Amendment_under_37_CFR_1_312_11_354569.pdf	110232 de7ace45e7b1c76fbe00187b315ac8ed0d3a8a77	no	7

### Warnings:

### Information:

JUNIPER Exhibit 1003

App. 3, pg. 12

'652 File History 012

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

**New Applications Under 35 U.S.C. 111**

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

**National Stage of an International Application under 35 U.S.C. 371**

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

**New International Application Filed with the USPTO as a Receiving Office**

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

**PART B - FEE(S) TRANSMITTAL**

Complete and send this form, together with applicable fee(s), to: **Mail Stop ISSUE FEE**  
**Commissioner for Patents**  
**P.O. Box 1450**  
**Alexandria, Virginia 22313-1450**  
 or **Fax (571)-273-2885**

**INSTRUCTIONS:** This form should be used for transmitting the **ISSUE FEE** and **PUBLICATION FEE** (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

**CURRENT CORRESPONDENCE ADDRESS** (Note: Use Block 1 for any change of address)

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

65638                      7590                      12/02/2010  
**OMIKRON IP LAW GROUP**  
 16325 Boones Ferry Rd.  
 SUITE 204  
 LAKE OSWEGO, OR 97035

**Certificate of Mailing or Transmission**  
 I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

<b>Lindsey Hunt</b>	(Depositor's name)
<i>Lindsey Hunt</i>	(Signature)
<b>March 1, 2011</b>	(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,369	02/14/2006	Pling Pan	002.P045	6912

TITLE OF INVENTION: PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	\$1510	\$0	\$0	\$1510	03/02/2011

EXAMINER	ART UNIT	CLASS-SUBCLASS
LIU, SIMING	2472	370-228000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).  
 Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.  
 "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47, Rev 03-02 or more recent) attached. Use of a Customer Number is required.
2. For printing on the patent front page, list  
 (1) the names of up to 3 registered patent attorneys or agents OR, alternatively, \_\_\_\_\_ 1  
 (2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed. \_\_\_\_\_ 2  
 \_\_\_\_\_ 3

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)  
 PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE: **Brixham Solutions LTD.**  
 (B) RESIDENCE: (CITY and STATE OR COUNTRY) **OMC Chambers, Wickhams Cay 1, Road Town Tortola, British Virgin Islands**

Please check the appropriate assignee category or categories (will not be printed on the patent):  Individual  Corporation or other private group entity  Government

- 4a. The following fee(s) are submitted:  
 Issue Fee  
 Publication Fee (No small entity discount permitted)  
 Advance Order - # of Copies \_\_\_\_\_
- 4b. Payment of Fee(s): (Please first reapply any previously paid issue fee shown above)  
 A check is enclosed.  
 Payment by credit card. Form PTO-2038 is attached.  
 The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number \_\_\_\_\_ (enclose an extra copy of this form).

5. Change in Entity Status (from status indicated above)  
 a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.  b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant, a registered attorney or agent, or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature: *Ted Crawford* Date: March 1, 2011  
 Typed or printed name: Ted Crawford Registration No.: 50,610

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, 115, Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.



Electronic Patent Application Fee Transmittal				
<b>Application Number:</b>	11354569			
<b>Filing Date:</b>	14-Feb-2006			
<b>Title of Invention:</b>	PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE			
<b>First Named Inventor/Applicant Name:</b>	Ping Pan			
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt			
<b>Attorney Docket Number:</b>	002.P045			
Filed as Large Entity				
<b>Utility under 35 USC 111(a) Filing Fees</b>				
Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
<b>Basic Filing:</b>				
<b>Pages:</b>				
<b>Claims:</b>				
<b>Miscellaneous-Filing:</b>				
<b>Petition:</b>				
<b>Patent-Appeals-and-Interference:</b>				
<b>Post-Allowance-and-Post-Issuance:</b>				
Utility Appl issue fee	1501	1	1510	1510
<b>Extension-of-Time:</b>				

Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
<b>Miscellaneous:</b>				
<b>Total in USD (\$)</b>				<b>1510</b>

## Electronic Acknowledgement Receipt

<b>EFS ID:</b>	9564085
<b>Application Number:</b>	11354569
<b>International Application Number:</b>	
<b>Confirmation Number:</b>	6912
<b>Title of Invention:</b>	PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE
<b>First Named Inventor/Applicant Name:</b>	Ping Pan
<b>Customer Number:</b>	65638
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt
<b>Filer Authorized By:</b>	Ted A. Crawford
<b>Attorney Docket Number:</b>	002.P045
<b>Receipt Date:</b>	01-MAR-2011
<b>Filing Date:</b>	14-FEB-2006
<b>Time Stamp:</b>	18:18:48
<b>Application Type:</b>	Utility under 35 USC 111(a)

### Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	\$1510
RAM confirmation Number	6272
Deposit Account	
Authorized User	

### File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip (if appl.)	Pages (if appl.)
-----------------	----------------------	-----------	-------------------------------------	-----------------------------	------------------

JUNIPER Exhibit 1003

App. 3, pg. 17

'652 File History 017

1	Issue Fee Payment (PTO-85B)	Issue_Fee_Transmittal_Signed_002_P045.pdf	928091 57dc3079e63010d7f14e01b515b9307acc0ea397	no	1
<b>Warnings:</b>					
<b>Information:</b>					
2	Fee Worksheet (PTO-875)	fee-info.pdf	30085 eca77e8ff2d6c47bb927f64983dce8a2838e43f6	no	2
<b>Warnings:</b>					
<b>Information:</b>					
<b>Total Files Size (in bytes):</b>				958176	
<p>This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.</p> <p><b><u>New Applications Under 35 U.S.C. 111</u></b>  If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.</p> <p><b><u>National Stage of an International Application under 35 U.S.C. 371</u></b>  If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.</p> <p><b><u>New International Application Filed with the USPTO as a Receiving Office</u></b>  If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.</p>					



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	002.P045	6912
65638	7590	12/22/2010	EXAMINER	
OMIKRON IP LAW GROUP 16325 Boones Ferry Rd. SUITE 204 LAKE OSWEGO, OR 97035			LIU, SIMING	
			ART UNIT	PAPER NUMBER
			2472	
			MAIL DATE	DELIVERY MODE
			12/22/2010	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>supplemental Notice of Allowability</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	11/354,569	PAN, PING	
	<b>Examiner</b>	<b>Art Unit</b>	
	SIMING LIU	2472	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--**

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1.  This communication is responsive to 12/06/2010.
2.  The allowed claim(s) is/are 1-5, 10-13, 17-18, 22-27.
3.  Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
  - a)  All    b)  Some\*    c)  None    of the:
    1.  Certified copies of the priority documents have been received.
    2.  Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
    3.  Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

\* Certified copies not received: \_\_\_\_\_.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.  
**THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.**

4.  A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
5.  CORRECTED DRAWINGS ( as "replacement sheets") must be submitted.
  - (a)  including changes required by the Notice of Draftsperson's Patent Drawing Review ( PTO-948) attached
    - 1)  hereto or 2)  to Paper No./Mail Date \_\_\_\_\_.
  - (b)  including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date \_\_\_\_\_.

**Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).**
6.  DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

**Attachment(s)**

- |  |  |
|--|--|
| 1. <input type="checkbox"/> Notice of References Cited (PTO-892)   | 5. <input type="checkbox"/> Notice of Informal Patent Application                      |
| 2. <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)   | 6. <input type="checkbox"/> Interview Summary (PTO-413),<br>Paper No./Mail Date _____. |
| 3. <input checked="" type="checkbox"/> Information Disclosure Statements (PTO/SB/08),<br>Paper No./Mail Date <u>12/06/2010</u> | 7. <input type="checkbox"/> Examiner's Amendment/Comment                               |
| 4. <input type="checkbox"/> Examiner's Comment Regarding Requirement for Deposit<br>of Biological Material                     | 8. <input type="checkbox"/> Examiner's Statement of Reasons for Allowance              |
|  | 9. <input type="checkbox"/> Other _____.   |

/S. L./  
Examiner, Art Unit 2472

/William Trost/  
Supervisory Patent Examiner, Art Unit 2472

Receipt date: 12/06/2010

11354569 - GALL: 2472

Doc code: IDS

Doc description: Information Disclosure Statement (IDS) Filed

Approved for use through 07/31/2012. OMB 0651-0031  
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Application Number		11354569
	Filing Date		2006-02-14
	First Named Inventor	Ping Pan	
	Art Unit	4145	
	Examiner Name	Liu, Siming	
	Attorney Docket Number	002.P045	

U.S. PATENTS <span style="float: right;">Remove</span>						
Examiner Initial*	Cite No	Patent Number	Kind Code <sup>1</sup>	Issue Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear
/S.L./	1	6985488		2006-01-10	Pan, et al.	
↓	2	6347088		2002-02-12	Katou, et al.	
	3	5920705		1999-07-06	Lyon et al.	
	4	6430184		2002-08-06	Robins, et al.	
	5	6813271		2004-11-02	Julian F. Cable	
	6	6621793		2004-11-02	Widegren et al.	
	7	6845389		2005-01-18	Sen et al.	
	↓	8	7050396		2006-05-23	Cohen et al.

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Receipt date: 12/06/2010		Application Number	11354569	11354569 - GAU: 2472	
			Filing Date	2006-02-14		
			First Named Inventor	Ping Pan		
			Art Unit	4145		
			Examiner Name	Liu, Siming		
			Attorney Docket Number	002.P045		

/S.L./	9	7436782		2008-10-14	NGO et al.	
	10	6546427		2003-04-07	Marni S. Ehrlich	
	11	6680943		2004-01-20	Gibson, et al.	
	12	6665273		2003-12-16	Goguen, et al.	
	13	6167051		2000-12-26	Nagami, et al.	
	14	6751684		2004-06-15	Owen, et al.	
	15	7697528		2007-05-03	Simon Parry	

If you wish to add additional U.S. Patent citation information please click the Add button.

**U.S.PATENT APPLICATION PUBLICATIONS**

Examiner Initial*	Cite No	Publication Number	Kind Code <sup>1</sup>	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear
/S.L./	1	20030039237	A1	2003-02-27	Jan E. Forslow	
/S.L./	2	20040105459	A1	2004-06-03	Raghu Mannam	



Receipt date: 12/06/2010

**INFORMATION DISCLOSURE  
STATEMENT BY APPLICANT**  
( Not for submission under 37 CFR 1.99)

Application Number	11354569	11354569 - GAU: 2472
Filing Date	2006-02-14	
First Named Inventor	Ping Pan	
Art Unit	4145	
Examiner Name	Liu, Siming	
Attorney Docket Number	002.P045	

/S.L./	3	20040174865		2004-09-09	Alan O'Neill	
	4	20040252717		2004-12-16	Solomon et al.	
	5	20050044262		2005-02-24	Wei Luo	
	6	20010023453	A1	2001-09-20	Jim Sundqvist	
	7	20020141393	A1	2002-10-03	Goran A.P. Eriksson	
	8	20040156313	A1	2004-08-12	Ralph Theodore Hofmeister	
	9	20050220148	A1	2005-10-06	Nick DelRegno	
	10	20060090008	A1	2006-04-27	Jim Guichard	
	11	20080031129	A1	2008-02-07	Jim Arseneault	
	12	20010021175		2001-09-13	Haverinen, Henry	
	13	20020146026		2002-10-10	Unitt, et al.	

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Receipt date: 12/06/2010		Application Number	11354569	11354569 - GAU: 2472	
			Filing Date	2006-02-14		
			First Named Inventor	Ping Pan		
			Art Unit	4145		
			Examiner Name	Liu, Siming		
			Attorney Docket Number	002.P045		

/S.L./	14	20070206607	A1	2007-09-06	John T. Chapman	
	15	20070127479	A1	2007-06-07	Sinicrope et al.	
	16	20070053366	A2	2007-03-08	Earl Hardin Booth III	
	17	20060233167	A1	2006-10-19	Shawn McAllister	
	18	20060146832	A1	2006-07-06	Sanjeev Rampal	
	19	20050237927	A1	2005-10-27	Kano, Shinya et al.	
	20	20040114595	A1	2004-06-17	Masami Doukai	
	21	20030002482		2003-01-01	Kubler, et al.	
	22	20020112072		2002-08-01	Sudhanshu Jain	
	23	20060002423	A1	2006-01-05	James William Rembert	
	24	20050018605	A1	2005-01-27	Richard Foote	

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Application Number	11354569	11354569 - GAU: 2472
	Filing Date	2006-02-14	
	First Named Inventor	Ping Pan	
	Art Unit	4145	
	Examiner Name	Liu, Siming	
	Attorney Docket Number	002.P045	

If you wish to add additional U.S. Published Application citation information please click the Add button.								<b>Add</b>
<b>FOREIGN PATENT DOCUMENTS</b>								<b>Remove</b>
Examiner Initial*	Cite No	Foreign Document Number <sup>3</sup>	Country Code <sup>2</sup> j	Kind Code <sup>4</sup>	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear	T <sup>5</sup>
	1							<input type="checkbox"/>
If you wish to add additional Foreign Patent Document citation information please click the Add button.								<b>Add</b>
<b>NON-PATENT LITERATURE DOCUMENTS</b>								<b>Remove</b>
Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.						T <sup>5</sup>
/S.L./	1	BRADEN, R. et al., "Integrated Services in the Internet Architecture: an overview," Network Working Group, June 1994						<input type="checkbox"/>
	2	BRYANT, S. et al., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture," Network Working Group, March 2005.						<input type="checkbox"/>
	3	BLAKE, S. et al., "An Architecture for Differentiated Services," Network Working Group, December 1998.						<input type="checkbox"/>
	4	SHAH, Himanshu et al., Internet Draft, ARP Mediation for IP Interworking of Layer 2 VPN, L2VPN Working Group, July 2007.						<input type="checkbox"/>
	5	MARTINI, Luca et al., Internet Draft, Segmented Pseudo Wire, Network Working Group, July 2007.						<input type="checkbox"/>
	6	PAN, P. et al., Internet Draft, Pseudo Wire Protection, July 2006.						<input type="checkbox"/>

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Application Number		11354569	11354569 - GAU: 2472
	Filing Date		2006-02-14	
	First Named Inventor	Ping Pan		
	Art Unit	4145		
	Examiner Name	Liu, Siming		
	Attorney Docket Number	002.P045		
	Receipt date: 12/06/2010			

/S.L./	7	ROSEN, Eric C. et al., Internet Draft, PWE3 Congestion Control Framework, Network Working Group, March 2004.	<input type="checkbox"/>
	8	ROSEN, E. et al., BGP-MPLS IP Virtual Private Networks (VPN), Network Working Group, February 2006.	<input type="checkbox"/>
	9	PAN, Ping, Internet Draft, Dry-Martini: Supporting Pseudo-wires in Sub-IP Access Networks, Network Working Group, July 2005.	<input type="checkbox"/>
	10	MCPHERSON et al., Pseudowire Emulation Edge to Edge (PWE3) June 13, 2007, <a href="http://www.ietf.org/html.charters/pwe3-carter.html">http://www.ietf.org/html.charters/pwe3-carter.html</a>	<input type="checkbox"/>
	11	AFFERTON, Thomas S. et al., Ethernet Transport over Wide Area Networks, Packet-Aware Transport for Metro Networks, IEEE Communications Magazine, pp. 120-127, March 2004.	<input type="checkbox"/>
	12	MARTINI, L. et al., Pseudowire Setup and Maintenance using the Label Distribution Protocol (LDP), Network Working Group, April 2006.	<input type="checkbox"/>
	13	ANDERSON, L. et al., LDP Specification, Network Working Group, January 2001.	<input type="checkbox"/>
	14	MARTINI, Luca et al., Encapsulation Methods for Transport of Ethernet over MPLS Networks, Network Working Group, April 2006.	<input type="checkbox"/>
	15	MARTINI, Luca et al., Encapsulation Methods for Transport of Frame Relay Over MPLS Networks, Network Working Group, February 2006.	<input type="checkbox"/>
	16	METZ, Chris et al., Pseudowire Attachment Identifiers for Aggregation and VPN Autodiscovery, PWE3 Working Group, February 25, 2006.	<input type="checkbox"/>
	17	MARTINI, Luca et al., Dynamic Placement of Multi Segment Pseudo Wires, PWE3 Working Group, June 2006.	<input type="checkbox"/>

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Application Number		11354569	11354569 - GAU: 2472
	Filing Date		2006-02-14	
	First Named Inventor	Ping Pan		
	Art Unit	4145		
	Examiner Name	Liu, Siming		
	Attorney Docket Number	002.P045		

/S.L./	18	MARTINI, Luca et al., "Pseudowire Setup and Maintenance using LDP", Network Working Group, March 2005.	<input type="checkbox"/>
/S.L./	19	VASSEUR, et al., Path Computation Element (pce), May 9, 2007, <a href="http://www.ietf.org/html.charters/pce.charter.html">http://www.ietf.org/html.charters/pce.charter.html</a>	<input type="checkbox"/>
/S.L./	20	THEIMER, T. et al, "Requirements for OAM Functionality in MPLS", October 1999, Watersprings.	<input type="checkbox"/>
/S.L./	21	Harry Newton, "Newton's Telecom Dictionary", 23rd Updated and Expanded Edition, p. 825,p. 239, Flatiron Publishing, New York, March 2007.	<input type="checkbox"/>

If you wish to add additional non-patent literature document citation information please click the Add button

**EXAMINER SIGNATURE**

Examiner Signature	/Siming Liu/	Date Considered	12/09/2010
--------------------	--------------	-----------------	------------

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> See Kind Codes of USPTO Patent Documents at [www.USPTO.GOV](http://www.USPTO.GOV) or MPEP 901.04. <sup>2</sup> Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). <sup>3</sup> For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. <sup>4</sup> Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. <sup>5</sup> Applicant is to place a check mark here if English language translation is attached.

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Application Number	11354569	11354569 - GAU: 2472
	Filing Date	2006-02-14	
	First Named Inventor	Ping Pan	
	Art Unit	4145	
	Examiner Name	Liu, Siming	
	Attorney Docket Number	002.P045	

**CERTIFICATION STATEMENT**

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

**OR**

That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

A certification statement is not submitted herewith.

**SIGNATURE**

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Ted A. Crawford/	Date (YYYY-MM-DD)	2010-12-06
Name/Print	Ted A. Crawford	Registration Number	50,610

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. **DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

**Privacy Act Statement**

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Application Number		11354569	
	Filing Date		2006-02-14	
	First Named Inventor	Ping Pan		
	Art Unit	4145		
	Examiner Name	Liu, Siming		
	Attorney Docket Number	002.P045		

U.S.PATENTS <span style="float: right;"><a href="#">Remove</a></span>						
Examiner Initial*	Cite No	Patent Number	Kind Code <sup>1</sup>	Issue Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear
	1	6985488		2006-01-10	Pan, et al.	
	2	6347088		2002-02-12	Katou, et al.	
	3	5920705		1999-07-06	Lyon et al.	
	4	6430184		2002-08-06	Robins, et al.	
	5	6813271		2004-11-02	Julian F. Cable	
	6	6621793		2004-11-02	Widegren et al.	
	7	6845389		2005-01-18	Sen et al.	
	8	7050396		2006-05-23	Cohen et al.	



**INFORMATION DISCLOSURE  
STATEMENT BY APPLICANT**  
( Not for submission under 37 CFR 1.99)

Application Number	11354569
Filing Date	2006-02-14
First Named Inventor	Ping Pan
Art Unit	4145
Examiner Name	Liu, Siming
Attorney Docket Number	002.P045

9	7436782		2008-10-14	NGO et al.	
10	6546427		2003-04-07	Marni S. Ehrlich	
11	6680943		2004-01-20	Gibson, et al.	
12	6665273		2003-12-16	Goguen, et al.	
13	6167051		2000-12-26	Nagami, et al.	
14	6751684		2004-06-15	Owen, et al.	
15	7697528		2007-05-03	Simon Parry	

If you wish to add additional U.S. Patent citation information please click the Add button.

Add

**U.S.PATENT APPLICATION PUBLICATIONS**

Remove

Examiner Initial*	Cite No	Publication Number	Kind Code <sup>1</sup>	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear
	1	20030039237	A1	2003-02-27	Jan E. Forslow	
	2	20040105459	A1	2004-06-03	Raghu Mannam	

**INFORMATION DISCLOSURE  
STATEMENT BY APPLICANT**  
( Not for submission under 37 CFR 1.99)

Application Number		11354569
Filing Date		2006-02-14
First Named Inventor	Ping Pan	
Art Unit	4145	
Examiner Name	Liu, Siming	
Attorney Docket Number	002.P045	

3	20040174865		2004-09-09	Alan O'Neill	
4	20040252717		2004-12-16	Solomon et al.	
5	20050044262		2005-02-24	Wei Luo	
6	20010023453	A1	2001-09-20	Jim Sundqvist	
7	20020141393	A1	2002-10-03	Goran A.P. Eriksson	
8	20040156313	A1	2004-08-12	Ralph Theodore Hofmeister	
9	20050220148	A1	2005-10-06	Nick DelRegno	
10	20060090008	A1	2006-04-27	Jim Guichard	
11	20080031129	A1	2008-02-07	Jim Arseneault	
12	20010021175		2001-09-13	Haverinen, Henry	
13	20020146026		2002-10-10	Unitt, et al.	

**INFORMATION DISCLOSURE  
STATEMENT BY APPLICANT**  
( Not for submission under 37 CFR 1.99)

Application Number		11354569
Filing Date		2006-02-14
First Named Inventor	Ping Pan	
Art Unit	4145	
Examiner Name	Liu, Siming	
Attorney Docket Number	002.P045	

	14	20070206607	A1	2007-09-06	John T. Chapman	
	15	20070127479	A1	2007-06-07	Sinicrope et al.	
	16	20070053366	A2	2007-03-08	Earl Hardin Booth III	
	17	20060233167	A1	2006-10-19	Shawn McAllister	
	18	20060146832	A1	2006-07-06	Sanjeev Rampal	
	19	20050237927	A1	2005-10-27	Kano, Shinya et al.	
	20	20040114595	A1	2004-06-17	Masami Doukai	
	21	20030002482		2003-01-01	Kubler, et al.	
	22	20020112072		2002-08-01	Sudhanshu Jain	
	23	20060002423	A1	2006-01-05	James William Rembert	
	24	20050018605	A1	2005-01-27	Richard Foote	

<b>INFORMATION DISCLOSURE STATEMENT BY APPLICANT</b> ( Not for submission under 37 CFR 1.99)	Application Number	11354569
	Filing Date	2006-02-14
	First Named Inventor	Ping Pan
	Art Unit	4145
	Examiner Name	Liu, Siming
	Attorney Docket Number	002.P045

If you wish to add additional U.S. Published Application citation information please click the Add button.

**FOREIGN PATENT DOCUMENTS**

Examiner Initial*	Cite No	Foreign Document Number <sup>3</sup>	Country Code <sup>2</sup> j	Kind Code <sup>4</sup>	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines where Relevant Passages or Relevant Figures Appear	T <sup>5</sup>
	1							<input type="checkbox"/>

If you wish to add additional Foreign Patent Document citation information please click the Add button.

**NON-PATENT LITERATURE DOCUMENTS**

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T <sup>5</sup>
	1	BRADEN, R. et al., "Integrated Services in the Internet Architecture: an overview," Network Working Group, June 1994	<input type="checkbox"/>
	2	BRYANT, S. et al., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture," Network Working Group, March 2005.	<input type="checkbox"/>
	3	BLAKE, S. et al., "An Architecture for Differentiated Services," Network Working Group, December 1998.	<input type="checkbox"/>
	4	SHAH, Himanshu et al., Internet Draft, ARP Mediation for IP Interworking of Layer 2 VPN, L2VPN Working Group, July 2007.	<input type="checkbox"/>
	5	MARTINI, Luca et al., Internet Draft, Segmented Pseudo Wire, Network Working Group, July 2007.	<input type="checkbox"/>
	6	PAN, P. et al., Internet Draft, Pseudo Wire Protection, July 2006.	<input type="checkbox"/>

**INFORMATION DISCLOSURE  
STATEMENT BY APPLICANT**  
( Not for submission under 37 CFR 1.99)

Application Number	11354569
Filing Date	2006-02-14
First Named Inventor	Ping Pan
Art Unit	4145
Examiner Name	Liu, Siming
Attorney Docket Number	002.P045

7	ROSEN, Eric C. et al., Internet Draft, PWE3 Congestion Control Framework, Network Working Group, March 2004.	<input type="checkbox"/>
8	ROSEN, E. et al., BGP-MPLS IP Virtual Private Networks (VPN), Network Working Group, February 2006.	<input type="checkbox"/>
9	PAN, Ping, Internet Draft, Dry-Martini: Supporting Pseudo-wires in Sub-IP Access Networks, Network Working Group, July 2005.	<input type="checkbox"/>
10	MCPHERSON et al., Pseudowire Emulation Edge to Edge (PWE3) June 13, 2007, <a href="http://www.ietf.org/html.charters/pwe3-carter.html">http://www.ietf.org/html.charters/pwe3-carter.html</a>	<input type="checkbox"/>
11	AFFERTON, Thomas S. et al., Ethernet Transport over Wide Area Networks, Packet-Aware Transport for Metro Networks, IEEE Communications Magazine, pp. 120-127, March 2004.	<input type="checkbox"/>
12	MARTINI, L. et al., Pseudowire Setup and Maintenance using the Label Distribution Protocol (LDP), Network Working Group, April 2006.	<input type="checkbox"/>
13	ANDERSON, L. et al., LDP Specification, Network Working Group, January 2001.	<input type="checkbox"/>
14	MARTINI, Luca et al., Encapsulation Methods for Transport of Ethernet over MPLS Networks, Network Working Group, April 2006.	<input type="checkbox"/>
15	MARTINI, Luca et al., Encapsulation Methods for Transport of Frame Relay Over MPLS Networks, Network Working Group, February 2006.	<input type="checkbox"/>
16	METZ, Chris et al., Pseudowire Attachment Identifiers for Aggregation and VPN Autodiscovery, PWE3 Working Group, February 25, 2006.	<input type="checkbox"/>
17	MARTINI, Luca et al., Dynamic Placement of Multi Segment Pseudo Wires, PWE3 Working Group, June 2006.	<input type="checkbox"/>

**INFORMATION DISCLOSURE  
STATEMENT BY APPLICANT**  
( Not for submission under 37 CFR 1.99)

Application Number	11354569
Filing Date	2006-02-14
First Named Inventor	Ping Pan
Art Unit	4145
Examiner Name	Liu, Siming
Attorney Docket Number	002.P045

18	MARTINI, Luca et al., "Pseudowire Setup and Maintenance using LDP", Network Working Group, March 2005.	<input type="checkbox"/>
19	VASSEUR, et al., Path Computation Element (pce), May 9, 2007, <a href="http://www.ietf.org/html.charters/pce.charter.html">http://www.ietf.org/html.charters/pce.charter.html</a>	<input type="checkbox"/>
20	THEIMER, T. et al, "Requirements for OAM Functionality in MPLS", October 1999, Watersprings.	<input type="checkbox"/>
21	Harry Newton, "Newton's Telecom Dictionary", 23rd Updated and Expanded Edition, p. 825,p. 239, Flatiron Publishing, New York, March 2007.	<input type="checkbox"/>

If you wish to add additional non-patent literature document citation information please click the Add button

**EXAMINER SIGNATURE**

Examiner Signature		Date Considered	
--------------------	--	-----------------	--

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup> See Kind Codes of USPTO Patent Documents at [www.USPTO.GOV](http://www.USPTO.GOV) or MPEP 901.04. <sup>2</sup> Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). <sup>3</sup> For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. <sup>4</sup> Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. <sup>5</sup> Applicant is to place a check mark here if English language translation is attached.

**INFORMATION DISCLOSURE  
STATEMENT BY APPLICANT**  
( Not for submission under 37 CFR 1.99)

Application Number	11354569
Filing Date	2006-02-14
First Named Inventor	Ping Pan
Art Unit	4145
Examiner Name	Liu, Siming
Attorney Docket Number	002.P045

**CERTIFICATION STATEMENT**

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

**OR**

That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

A certification statement is not submitted herewith.

**SIGNATURE**

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Ted A. Crawford/	Date (YYYY-MM-DD)	2010-12-06
Name/Print	Ted A. Crawford	Registration Number	50,610

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. **DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

## Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.



Electronic Patent Application Fee Transmittal				
<b>Application Number:</b>	11354569			
<b>Filing Date:</b>	14-Feb-2006			
<b>Title of Invention:</b>	PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE			
<b>First Named Inventor/Applicant Name:</b>	Ping Pan			
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt			
<b>Attorney Docket Number:</b>	002.P045			
Filed as Large Entity				
<b>Utility under 35 USC 111(a) Filing Fees</b>				
Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
<b>Basic Filing:</b>				
<b>Pages:</b>				
<b>Claims:</b>				
<b>Miscellaneous-Filing:</b>				
<b>Petition:</b>				
<b>Patent-Appeals-and-Interference:</b>				
<b>Post-Allowance-and-Post-Issuance:</b>				
<b>Extension-of-Time:</b>				

Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
<b>Miscellaneous:</b>				
Submission- Information Disclosure Stmt	1806	1	180	180
<b>Total in USD (\$)</b>				<b>180</b>

## Electronic Acknowledgement Receipt

<b>EFS ID:</b>	8972326
<b>Application Number:</b>	11354569
<b>International Application Number:</b>	
<b>Confirmation Number:</b>	6912
<b>Title of Invention:</b>	PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE
<b>First Named Inventor/Applicant Name:</b>	Ping Pan
<b>Customer Number:</b>	65638
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt
<b>Filer Authorized By:</b>	Ted A. Crawford
<b>Attorney Docket Number:</b>	002.P045
<b>Receipt Date:</b>	06-DEC-2010
<b>Filing Date:</b>	14-FEB-2006
<b>Time Stamp:</b>	14:40:33
<b>Application Type:</b>	Utility under 35 USC 111(a)

### Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	\$ 180
RAM confirmation Number	853
Deposit Account	
Authorized User	

### File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip (if appl.)	Pages (if appl.)
-----------------	----------------------	-----------	-------------------------------------	-----------------------------	------------------

JUNIPER Exhibit 1003

App. 3, pg. 41

'652 File History 041

1	NPL Documents	NPL_1.pdf	89990	no	31
			a6127715a854785c51ca36d74fecac196ce c3e7		
<b>Warnings:</b>					
<b>Information:</b>					
2	NPL Documents	NPL_2.pdf	63198	no	42
			14e78a80f02a093f77d6c0f7278a4834e60b febb		
<b>Warnings:</b>					
<b>Information:</b>					
3	NPL Documents	NPL_3.pdf	96518	no	34
			f24be07c4479551fa6f5951b528edf742a5c d442		
<b>Warnings:</b>					
<b>Information:</b>					
4	NPL Documents	NPL_4.pdf	140721	no	26
			8089edf7106301fdb052c69926628d5dc77 e30ad		
<b>Warnings:</b>					
<b>Information:</b>					
5	NPL Documents	NPL_5.pdf	1365045	no	36
			de0236b49cfd6b47544ad59471179b3e343 cd1de		
<b>Warnings:</b>					
<b>Information:</b>					
6	NPL Documents	NPL_6.pdf	68343	no	18
			27e4e2d324d364072bffc8cf33b6e096a58 33635		
<b>Warnings:</b>					
<b>Information:</b>					
7	NPL Documents	NPL_7.pdf	53753	no	26
			e83d118bd079b81749b7af6e990ae093424 e2487		
<b>Warnings:</b>					
<b>Information:</b>					
8	NPL Documents	NPL_8.pdf	2052028	no	44
			62adbcd339ac7cad39761c6d0535ca7155 d7cee		
<b>Warnings:</b>					
<b>Information:</b>					
9	NPL Documents	NPL_9.pdf	60587	no	35
			2c53c1260d216d9c46429f888119287c40c 99d3		
<b>Warnings:</b>					
<b>Information:</b>					

10	NPL Documents	NPL_10.pdf	251348	no	5
			60b59e1b9482bbb52f92fd9d3efd27712e3d2de4		
<b>Warnings:</b>					
<b>Information:</b>					
11	NPL Documents	NPL_11.pdf	1399249	no	8
			9a83a1bdf1b789b74956254e63f8dd701c0ed4a		
<b>Warnings:</b>					
<b>Information:</b>					
12	NPL Documents	NPL_12.pdf	1218073	no	30
			66b68bf2e62dc3aab0ef7804f26acea08b4b7c28		
<b>Warnings:</b>					
<b>Information:</b>					
13	NPL Documents	NPL_13.pdf	4386129	no	120
			150957e1d6ecb210f509ab2a72139d335c43e71d		
<b>Warnings:</b>					
<b>Information:</b>					
14	NPL Documents	NPL_14.pdf	772229	no	22
			8c88719ae84466d5805ac126bd3bfe09362d231e		
<b>Warnings:</b>					
<b>Information:</b>					
15	NPL Documents	NPL_15.pdf	650941	no	19
			938954e9b58c23f70f649baa7fee1191ec61b99		
<b>Warnings:</b>					
<b>Information:</b>					
16	NPL Documents	NPL_16.pdf	278396	no	8
			d29fdb9de5378bd0f6a8d9d6c5c32a79a737e36		
<b>Warnings:</b>					
<b>Information:</b>					
17	NPL Documents	NPL_17.pdf	628969	no	19
			761d9e465812fb755bcbcb0ed9cccd4320af2de62		
<b>Warnings:</b>					
<b>Information:</b>					
18	NPL Documents	NPL_18.pdf	1126682	no	33
			5428174f9f6c07371a8f8a03b32f5eb4ac06a483		
<b>Warnings:</b>					
<b>Information:</b>					

19	NPL Documents	NPL_19.pdf	199537 8cae6ebd31835020465d93b0318782c5580db6b9	no	4
<b>Warnings:</b>					
<b>Information:</b>					
20	NPL Documents	NPL_20.pdf	198340 0395cd8fb6824ed413e679546c1d722323b4ab0a	no	5
<b>Warnings:</b>					
<b>Information:</b>					
21	NPL Documents	NPL_21.pdf	866367 43b4f28e3ced59540309d1c1d9ddb6701c95b79a	no	2
<b>Warnings:</b>					
<b>Information:</b>					
22	Information Disclosure Statement (IDS) Filed (SB/08)	IDS_002_P045.pdf	614892 ffc834a82c1e3defebbb58edbb75de31656db28fb	no	9
<b>Warnings:</b>					
<b>Information:</b>					
23	Fee Worksheet (PTO-875)	fee-info.pdf	30440 8861849ee10105e9df1ddbe891c71f55b8cadd64	no	2
<b>Warnings:</b>					
<b>Information:</b>					
<b>Total Files Size (in bytes):</b>				16611775	
<p>This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.</p> <p><b><u>New Applications Under 35 U.S.C. 111</u></b>  If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.</p> <p><b><u>National Stage of an International Application under 35 U.S.C. 371</u></b>  If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.</p> <p><b><u>New International Application Filed with the USPTO as a Receiving Office</u></b>  If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.</p>					



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P. O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

65638 7590 12/02/2010
OMIKRON IP LAW GROUP
16325 Boones Ferry Rd.
SUITE 204
LAKE OSWEGO, OR 97035

EXAMINER
LIU, SIMING
ART UNIT PAPER NUMBER
2472
DATE MAILED: 12/02/2010

Table with 5 columns: APPLICATION NO., FILING DATE, FIRST NAMED INVENTOR, ATTORNEY DOCKET NO., CONFIRMATION NO.
11/354,569 02/14/2006 Ping Pan 002.P045 6912

TITLE OF INVENTION: PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE

Table with 7 columns: APPLN. TYPE, SMALL ENTITY, ISSUE FEE DUE, PUBLICATION FEE DUE, PREV. PAID ISSUE FEE, TOTAL FEE(S) DUE, DATE DUE
nonprovisional NO \$1510 \$0 \$0 \$1510 03/02/2011

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the SMALL ENTITY status shown above.

If the SMALL ENTITY is shown as YES, verify your current SMALL ENTITY status:

- A. If the status is the same, pay the TOTAL FEE(S) DUE shown above.
B. If the status above is to be removed, check box 5b on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and twice the amount of the ISSUE FEE shown above, or

If the SMALL ENTITY is shown as NO:

- A. Pay TOTAL FEE(S) DUE shown above, or
B. If applicant claimed SMALL ENTITY status before, or is now claiming SMALL ENTITY status, check box 5a on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and 1/2 the ISSUE FEE shown above.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

**PART B - FEE(S) TRANSMITTAL**

**Complete and send this form, together with applicable fee(s), to: Mail Mail Stop ISSUE FEE  
 Commissioner for Patents  
 P.O. Box 1450  
 Alexandria, Virginia 22313-1450  
 or Fax (571)-273-2885**

**INSTRUCTIONS:** This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

65638                      7590                      12/02/2010  
**OMIKRON IP LAW GROUP**  
 16325 Boones Ferry Rd.  
 SUITE 204  
 LAKE OSWEGO, OR 97035

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

**Certificate of Mailing or Transmission**

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

_____ (Depositor's name)
_____ (Signature)
_____ (Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	002.P045	6912

TITLE OF INVENTION: PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	\$1510	\$0	\$0	\$1510	03/02/2011

EXAMINER	ART UNIT	CLASS-SUBCLASS
LIU, SIMING	2472	370-228000

<p>1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).  <input type="checkbox"/> Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.  <input type="checkbox"/> "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. <b>Use of a Customer Number is required.</b></p>	<p>2. For printing on the patent front page, list                  (1) the names of up to 3 registered patent attorneys or agents OR, alternatively, _____ 1                  (2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed. _____ 2                  _____ 3</p>
---	---

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)  
 PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.  
 (A) NAME OF ASSIGNEE \_\_\_\_\_ (B) RESIDENCE: (CITY and STATE OR COUNTRY) \_\_\_\_\_

Please check the appropriate assignee category or categories (will not be printed on the patent) :  Individual  Corporation or other private group entity  Government

<p>4a. The following fee(s) are submitted:  <input type="checkbox"/> Issue Fee  <input type="checkbox"/> Publication Fee (No small entity discount permitted)  <input type="checkbox"/> Advance Order - # of Copies _____</p>	<p>4b. Payment of Fee(s): (<b>Please first reapply any previously paid issue fee shown above</b>)  <input type="checkbox"/> A check is enclosed.  <input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.  <input type="checkbox"/> The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number _____ (enclose an extra copy of this form).</p>
---	---

5. **Change in Entity Status** (from status indicated above)  
 a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.       b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature \_\_\_\_\_ Date \_\_\_\_\_  
 Typed or printed name \_\_\_\_\_ Registration No. \_\_\_\_\_

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.





UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P. O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

Table with 5 columns: APPLICATION NO., FILING DATE, FIRST NAMED INVENTOR, ATTORNEY DOCKET NO., CONFIRMATION NO. Includes application details for Ping Pan and OMIKRON IP LAW GROUP.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment to date is 582 day(s). If the issue fee is paid on the date that is three months after the mailing date of this notice and the patent issues on the Tuesday before the date that is 28 weeks (six and a half months) after the mailing date of this notice, the Patent Term Adjustment will be 582 day(s).

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

<b>Notice of Allowability</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	11/354,569	PAN, PING	
	<b>Examiner</b>	<b>Art Unit</b>	
	SIMING LIU	2472	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--**

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1.  This communication is responsive to 10/21/2010.
2.  The allowed claim(s) is/are 1-5, 10-13, 17-18, 22-27.
3.  Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
  - a)  All    b)  Some\*    c)  None    of the:
    1.  Certified copies of the priority documents have been received.
    2.  Certified copies of the priority documents have been received in Application No. \_\_\_\_ .
    3.  Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

\* Certified copies not received: \_\_\_\_.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.

**THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.**

4.  A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
  5.  CORRECTED DRAWINGS ( as "replacement sheets") must be submitted.
    - (a)  including changes required by the Notice of Draftsperson's Patent Drawing Review ( PTO-948) attached
      - 1)  hereto or 2)  to Paper No./Mail Date \_\_\_\_.
    - (b)  including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date \_\_\_\_.
- Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).**
6.  DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

**Attachment(s)**

- |  |   |
|--|---|
| <ol style="list-style-type: none"> <li>1. <input type="checkbox"/> Notice of References Cited (PTO-892)</li> <li>2. <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)</li> <li>3. <input type="checkbox"/> Information Disclosure Statements (PTO/SB/08),<br/>Paper No./Mail Date ____</li> <li>4. <input type="checkbox"/> Examiner's Comment Regarding Requirement for Deposit of Biological Material</li> </ol> | <ol style="list-style-type: none"> <li>5. <input type="checkbox"/> Notice of Informal Patent Application</li> <li>6. <input type="checkbox"/> Interview Summary (PTO-413),<br/>Paper No./Mail Date ____ .</li> <li>7. <input type="checkbox"/> Examiner's Amendment/Comment</li> <li>8. <input checked="" type="checkbox"/> Examiner's Statement of Reasons for Allowance</li> <li>9. <input type="checkbox"/> Other ____.</li> </ol> |
|--|---|

/S. L./  
Examiner, Art Unit 2472

/William Trost/  
Supervisory Patent Examiner, Art Unit 2472

## DETAILED ACTION

### *Allowable Subject Matter*

1. Claims 1-5, 10-13, 17-18, 22-27 are allowed.
2. The following is an examiner's statement of reasons for allowance: with respect to claims 1, 11, 17, in addition to other limitations in the claims, the prior art fails to teach or disclose the specific of applicant's invention as claimed, particularly the feature describing "accepting the Pseudowire protection configuration parameter by the destination node; using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter; and determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based on, at least in part, on the priority for the standby Pseudowire".
3. Dependent claims 2-45, 10, 12-13, 18, 22-27 are allowable by virtue of their dependencies.
4. Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

**Conclusion**

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIMING LIU whose telephone number is (571)270-3859. The examiner can normally be reached on Monday-Friday 8:30am-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, William Trost can be reached on 571-272-7872. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/S. L./  
Examiner, Art Unit 2416

/William Trost/  
Supervisory Patent Examiner, Art  
Unit 2472

## EAST Search History

## EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	299	pseudowire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L2	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L6	79	(pseudowire or pseudo-wire) and initi\$5 and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L7	3083	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L8	27	(370/216,225,228).ccls. and L2	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L9	9	(709/220).ccls. and L2	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L10	41	((PING) near2 (PAN)).INV.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L11	2	((PING) near2 (PAN)).INV. and pseudowire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L12	2	((PING) near2 (PAN)).INV. and (pseudowire).clm.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L16	248	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L17	72	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L18	18	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (standby or backup)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L19	16	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (primary or main) and (secondly or backup or standby)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L20	99	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16

L21	31	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L22	44	(pseudowire or (pseudo wire) or pseudo-wire) and ((config\$7) same (destination near5 node))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L23	5	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with acknowledgement)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L24	17	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L25	529	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L26	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:16
L27	627	pseudowire or pseudo-wire or (pseudo wire)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L28	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (L27) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L29	255	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:16
L30	4093995	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2010/11/15 14:16
L31	1615408	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L32	4357729	(alter\$7 or backup or standby)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L33	29928696	@ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16

L34	7473513	(pick\$5 or select\$5 or choos\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L35	5	(restoration scheme) and (priority) and (standby mode)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L36	4454154	(send\$7 or transmit\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L37	784	(source) with L36 with (parameter\$1) with (destin\$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L38	67	(source node) with L36 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L39	4127	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L40	0	(ack or acknowledgement) same (config\$7 parameter\$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L41	29622	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L42	56520	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L43	33	(ack or acknowledgement) and (restoration scheme) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L45	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16

L46	11073	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L47	813	handshaking and @ad<"20050214" and (L30 with L31)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L48	119	((virtual path) and ((protection or restoration) near5 scheme) and priority	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L49	111	((virtual path) and ((protection or restoration) near5 scheme) and priority and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L50	4555	((protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L51	2755	((protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L52	7	((protection or restoration) near5 parameter) with (L36) and (destin\$7 near3 node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L53	28	((protection or restoration) near5 parameter) and (handshaking) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L54	79	((protection or restoration) near5 parameter) and (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L55	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L56	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16



L57	18	((protection or restoration or config\$7) near5 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L58	6	((protection or restoration or config\$7) near2 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L59	714	handshaking and @ad<"20050214" and (config\$7 parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L60	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16
L61	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16
L62	369	receiv\$7 acknowledgement destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16
L63	5	receiv\$7 acknowledgement accept\$3 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16
L64	11	receiv\$7 acknowledgement parameter accept\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16
L65	1164	(domain type) with (parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16
L66	4	(parameter) near5 includ\$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16
L67	4	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16

L68	75	(parameter\$1) with (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:16
L69	0	(domain type) with (single-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L70	0	(domain type) with (single near5 hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L71	75	(domain type) with parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L72	342	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L73	18	(single-hop) same (multi-hop) same (parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L74	233	field with indica\$7 with (topology)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L75	11	field with indica\$7 with (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L76	342	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L77	27	(field or parameter) same ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L78	538	((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16

L79	50	(parameter or field) same L78	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L80	0	L78 same (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L81	157	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L82	0	parameter with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L83	0	(field or parameter) with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L84	0	(field or parameter) with (show\$3 or indicat\$5) same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L85	0	(field or parameter) with (domain type) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:16
L86	72	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L87	16	(field or parameter) with (indicat\$5 or show\$5) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L88	2	(protection type) and (standby path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17

L89	3319	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L90	295	(hot and cold) same standby and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L91	52	(hot and cold) and (parameter with standby) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L92	21	(field with indicat\$5 with (standby mode)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L93	739	config\$9 with (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L94	425	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L95	339	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L96	204	config\$9 near3 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L97	7	config\$9 near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L98	10	type with (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L99	4	type near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17

L100	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L101	157	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L102	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L103	149	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L104	16	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) same (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L105	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 netowrk)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L106	249	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 network)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L107	0	(paramete or field or bit) with indicat\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L108	21	(paramete or field or bit) with indicat\$7 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L109	1	(protection near3 properties) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2010/11/15 14:17
L110	8	(protection near3 (parameter or propert\$5)) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2010/11/15 14:17
L111	82047	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/11/15 14:17

L112	19	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/11/15 14:17
L113	52	(pseudowire or pseudo-wire) and (standby)	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L114	7	(pseudowire or pseudo-wire) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L115	13	(protection scheme) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L116	1	(protection scheme) with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L117	1	(protection (type or property)) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L118	12	((protection (type or property)) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L119	23	(( (type or property) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L120	1	(protection scheme) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L121	2	(protection scheme) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L122	3	(protection (scheme or propert\$3 or parameter or type)) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L123	24	(backup path) with (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L124	0	(backup path) with (protection near3 parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L125	155	(backup path) with (protection )	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L126	1	(genera\$5 or configur\$5) with (backup path) with (parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L127	197	(genera\$5 or configur\$5) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L128	27	(genera\$5 or configur\$5) with (backup path) and (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17

L129	216	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L130	19	(genera\$5 or configur\$5 or setup) with (backup path) not L127	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L131	3083	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L132	4903	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L133	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L134	9	(709/220).ccls. and L133	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L137	662	(pseudowire or pseudo-wire or pseudo wire)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L138	22	(pseudowire or pseudo-wire or pseudo wire) and (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L139	299	pseudowire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L140	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L142	4	L140 with protection and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L145	300	(pseudowire or pseudo-wire) and initi\$5	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L146	79	(pseudowire or pseudo-wire) and initi\$5 and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L147	3083	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L148	27	(370/216,225,228).ccls. and L140	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L149	9	(709/220).ccls. and L140	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L150	41	((PING) near2 (PAN)).INV.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L151	2	((PING) near2 (PAN)).INV. and pseudowire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L152	2	((PING) near2 (PAN)).INV. and (pseudowire).clm.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L153	147	L140 and (primary)	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L154	36	L140 and (primary) and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L155	120	L140 and (config\$7) and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17

L156	14	TDM pseudowire	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L157	248	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L158	72	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad- "20050214"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L159	18	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad- "20050214" and (standby or backup)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L160	16	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad- "20050214" and (primary or main) and (secondly or backup or standby)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L161	99	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L162	31	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L163	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L164	44	(pseudowire or (pseudo wire) or pseudo-wire) and ((config\$7) same (destination near5 node))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L165	5	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with acknowledgement)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L166	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L167	17	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L168	529	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledgement))	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17



L169	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (L165) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L170	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L171	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (L170) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L172	627	pseudowire or pseudo-wire or (pseudo wire)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L173	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (L172) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L174	255	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L176	4093995	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2010/11/15 14:17
L177	1615408	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L179	40	(restoration scheme) and ("1:N")	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L180	5	(restoration scheme) and (priority) and (standby mode)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L181	29	(restoration scheme) and (priority) and (config\$7 near\$5 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L182	4127	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17

L183	0	(ack or acknowledgement) same (config\$7 parameter \$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L184	29622	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L185	56520	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L186	33	(ack or acknowledgement) and (restoration scheme) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L187	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L188	11073	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L189	813	handshaking and @ad<"20050214" and (L176 with L177)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L191	119	(virtual path) and ((protection or restoration) near5 scheme) and priority	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L192	111	(virtual path) and ((protection or restoration) near5 scheme) and priority and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L193	4555	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L194	2755	(protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17

L195	28	((protection or restoration) near5 parameter) and (handshaking) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L196	79	((protection or restoration) near5 parameter) and (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L197	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L198	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L199	18	((protection or restoration or config\$7) near5 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L200	6	((protection or restoration or config\$7) near2 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L201	714	handshaking and @ad<"20050214" and (config\$7 parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L202	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L203	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L204	369	receiv\$7 acknowledgement destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L205	0	receiv\$7 acknowledgement (parameter accept\$5 destination node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17

L206	5	receiv\$7 acknowledgement accept\$3 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L207	11	receiv\$7 acknowledgement parameter accept\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L208	2	"20030117950".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L209	1164	(domain type) with (parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L210	4	(parameter) near5 includ\$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L211	4	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L212	75	(parameter\$1) with (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:17
L213	75	(domain type) with parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L214	342	(single-hop) same (multi- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L215	18	(single-hop) same (multi- hop) same (parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L216	0	field with indica\$7 with ((single-hop) same (multi- hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17

L217	233	field with indica\$7 with (topology)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L218	11	field with indica\$7 with (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L219	342	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L220	27	(field or parameter) same ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L221	538	((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L222	50	(parameter or field) same L221	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L223	157	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L224	0	(field or parameter) with (domain type) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L225	72	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L226	16	(field or parameter) with (indicat\$5 or show\$5) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L227	2	(protection type) and (standby path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17

L228	3319	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L229	295	(hot and cold) same standby and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L230	52	(hot and cold) and (parameter with standby) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L231	21	(field with indicat\$5 with (standby mode)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L232	739	config\$9 with (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L233	425	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L234	339	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L235	204	config\$9 near3 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L236	7	config\$9 near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L237	10	type with (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L238	4	type near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17

L239	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L240	157	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L241	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L242	149	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L243	16	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) same (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L245	249	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 network)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L246	0	(parameter or field or bit) with indicator\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L247	21	(parameter or field or bit) with indicator\$7 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:17
L248	1	(protection near3 properties) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2010/11/15 14:17
L249	8	(protection near3 (parameter or property\$5)) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2010/11/15 14:17
L250	82047	config\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/11/15 14:17
L251	19	config\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/11/15 14:17
L252	52	(pseudowire or pseudo-wire) and (standby)	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17

L253	7	(pseudowire or pseudo-wire) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L254	13	(protection scheme) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L255	1	(protection scheme) with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L256	1	(protection (type or property)) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L257	12	((protection (type or property)) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L258	23	(( (type or property) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L259	1	(protection scheme) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L260	2	(protection scheme) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L261	3	(protection (scheme or property) or parameter or type) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L262	24	(backup path) with (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L263	0	(backup path) with (protection near parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L264	155	(backup path) with (protection )	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L265	1	(general or configuration) with (backup path) with (parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L266	197	(general or configuration) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L267	27	(general or configuration) with (backup path) and (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L268	0	(general or configuration) with (backup path) with (base or according)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L269	0	(general or configuration) with (backup path) with ("base" or "according")	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17



L270	216	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L271	19	(genera\$5 or configur\$5 or setup) with (backup path) not L266	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L272	3083	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L273	4903	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L274	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L275	9	(709/220).ccls. and L274	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L276	1	"20050226215"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/11/15 14:17
L277	2	"20060045028"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/11/15 14:17
L278	662	(pseudowire or pseudo-wire or pseudo wire)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L279	22	(pseudowire or pseudo-wire or pseudo wire) and (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:17
L280	106738	standby path with priority	US-PGPUB; USPAT	OR	ON	2010/11/15 14:17
L282	3	standby path with priority	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L283	7	switch\$5 with priority same (standby path)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L284	1	preempt with existing traffic	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L285	214	preempt with traffic	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L286	70	preempt with traffic with priority	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L287	67	LDp same acknow\$11	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L288	56	LDp same acknow\$11 and @ad<"20050216"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L289	12	LDp same acknow\$11 and @ad<"20050216" and (label distribution protocol)	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L290	0	Ping near2 pan and pseduo \$5	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L291	0	(Ping near2 pan).inv. and pseduo\$5	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17
L292	7	(Ping near2 pan).inv. and pseudo\$5	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:17

L293	199	((pseudo-wire) or (pseudowire)) and LDP	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:18
L294	60	((pseudo-wire) or (pseudowire)) and LDP and @ad<"20050216"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:18
L295	2	L294 and backup path	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:18
L296	44	LDP with protection	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:18
L297	33	LDP with protection and @ad<"20050216"	US-PGPUB; USPAT	ADJ	ON	2010/11/15 14:18
L298	8	protection with domain type	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/11/15 14:18
L299	5	protection and domain type and LDP	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/11/15 14:18
L300	65	single-hop and multi-hop and backup	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/11/15 14:18
L301	3	single-hop and multi-hop and backup path	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/11/15 14:18
L302	7	multi-hop pseudowire	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/11/15 14:18
L303	33	(\$7hop) with pseudowire	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:18
L304	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:18
L305	4127	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L306	29622	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L307	56520	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L308	11073	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18

L309	4555	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L310	2755	(protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L311	342	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L312	72	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L313	3319	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L314	425	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L315	339	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L316	82047	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/11/15 14:18
L317	216	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:18
L318	3083	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:18
L319	4903	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:18
L320	662	(pseudowire or pseudo-wire or pseudo wire)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:18
L321	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:18
L322	507	backhaul connection	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/11/15 14:18

L324	188	pseudowire same mpl	US-PGPUB; USPAT	OR	ON	2010/11/15 14:18
L325	497	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/11/15 14:18
L326	4127	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L327	29622	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L328	56520	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L329	11073	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L330	4555	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L331	2755	(protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L332	4	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/11/15 14:18
L333	342	(single-hop) same (multi- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L334	72	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L335	425	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18

L336	339	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/11/15 14:18
L337	216	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/11/15 14:18
L338	3083	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:18
L339	4903	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2010/11/15 14:18

**EAST Search History (Interference)**

<This search history is empty>

**11 / 15 / 2010 4:15:46 PM**


**C:\ Documents and Settings\ sliu3\ My Documents\ EAST\ Workspaces\ 11354569.wsp**

<b>Index of Claims</b>  	<b>Application/Control No.</b> 11354569	<b>Applicant(s)/Patent Under Reexamination</b> PAN, PING
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 2472

✓	<b>Rejected</b>	-	<b>Cancelled</b>	N	<b>Non-Elected</b>	A	<b>Appeal</b>
=	<b>Allowed</b>	÷	<b>Restricted</b>	I	<b>Interference</b>	O	<b>Objected</b>

Claims renumbered in the same order as presented by applicant
  CPA
  T.D.
  R.1.47

CLAIM		DATE								
Final	Original	10/30/2008	06/17/2009	01/14/2010	07/16/2010	11/29/2010				
1	1	✓	✓	✓	✓	=				
2	2	✓	✓	✓	✓	=				
3	3	✓	✓	✓	✓	=				
4	4	✓	✓	✓	✓	=				
5	5	✓	✓	✓	✓	=				
	6	✓	✓	✓	-	-				
	7	✓	✓	✓	-	-				
	8	✓	✓	✓	-	-				
	9	✓	✓	✓	-	-				
6	10	✓	✓	✓	✓	=				
9	11	✓	✓	✓	✓	=				
10	12	✓	✓	✓	✓	=				
11	13	✓	✓	✓	✓	=				
	14	✓	✓	✓	-	-				
	15	✓	✓	✓	-	-				
	16	✓	✓	✓	-	-				
14	17	✓	✓	✓	✓	=				
15	18	✓	✓	✓	✓	=				
	19	✓	✓	✓	-	-				
	20	✓	✓	✓	-	-				
	21	✓	✓	✓	-	-				
7	22				✓	=				
8	23				✓	=				
12	24				✓	=				
13	25				✓	=				
16	26				✓	=				
17	27				✓	=				


<b>Search Notes</b>  	<b>Application/Control No.</b>  11354569	<b>Applicant(s)/Patent Under Reexamination</b>  PAN, PING
	<b>Examiner</b>  SIMING LIU	<b>Art Unit</b>  2472

SEARCHED			
Class	Subclass	Date	Examiner
370	216, 225, 228	10/30/2008	/SL/
709	220	10/30/2008	/SL/
above	update search	6/17/2009	/SL/
update search	ABOVE	1/14/2010	/SL/
update search	ABOVE	7/16/2010	/SL/
update search	Above	11/22/2010	/SL/

SEARCH NOTES		
Search Notes	Date	Examiner
East Class search	11/10/2008 update 6/17/2009	/SL/
Palm inventor name search	10/30/2008 update 6/17/2009	/SL/
Consulted 101 issues with Peng, John	11/10/2008	/SL/
update search: ABOVE	1/14/2010	/SL/
update search: ABOVE	7/16/2010	/SL/
update search: ABOVE	11/22/2010	/SL/

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner
370	216, 225, 228	11/22/2010	/SL/

--	--

<b>Issue Classification</b> 	<b>Application/Control No.</b> 11354569	<b>Applicant(s)/Patent Under Reexamination</b> PAN, PING
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 2472

ORIGINAL					INTERNATIONAL CLASSIFICATION														
CLASS		SUBCLASS			CLAIMED					NON-CLAIMED									
370		228			H	0	4	J	3 / 14 (2006.01.01)										
<b>CROSS REFERENCE(S)</b>																			
CLASS	SUBCLASS (ONE SUBCLASS PER BLOCK)																		
370	216	225																	
709	220																		

<input type="checkbox"/> Claims renumbered in the same order as presented by applicant <input type="checkbox"/> CPA <input type="checkbox"/> T.D. <input type="checkbox"/> R.1.47															
Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	1	14	17												
2	2	15	18												
3	3		19												
4	4		20												
5	5		21												
	6	7	22												
	7	8	23												
	8	12	24												
	9	13	25												
6	10	16	26												
9	11	17	27												
10	12														
11	13														
	14														
	15														
	16														

/S. L./ Examiner.Art Unit 2472  (Assistant Examiner)	11/29/2010  (Date)	<b>Total Claims Allowed:</b> 17	
/WILLIAM TROST IV/ Supervisory Patent Examiner.Art Unit 2472  (Primary Examiner)	11/30/2010  (Date)	O.G. Print Claim(s) 1	O.G. Print Figure 5





UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	002.P045	6912
65638	7590	11/17/2010	EXAMINER	
OMIKRON IP LAW GROUP 16325 Boones Ferry Rd. SUITE 204 LAKE OSWEGO, OR 97035			LIU, SIMING	
			ART UNIT	PAPER NUMBER
			2472	
			MAIL DATE	DELIVERY MODE
			11/17/2010	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Notice of Panel Decision from Pre-Appeal Brief Review</b>	<b>Application/Control No.</b>		<b>Applicant(s)/Patent under Reexamination</b>	
	11/354,569		PAN, PING	
	WILLIAM TROST IV		<b>Art Unit</b>	
		2472		

This is in response to the Pre-Appeal Brief Request for Review filed 21 October 2010.

1.  **Improper Request** – The Request is improper and a conference will not be held for the following reason(s):

- The Notice of Appeal has not been filed concurrent with the Pre-Appeal Brief Request.
- The request does not include reasons why a review is appropriate.
- A proposed amendment is included with the Pre-Appeal Brief request.
- Other: .

The time period for filing a response continues to run from the receipt date of the Notice of Appeal or from the mail date of the last Office communication, if no Notice of Appeal has been received.

2.  **Proceed to Board of Patent Appeals and Interferences** – A Pre-Appeal Brief conference has been held. The application remains under appeal because there is at least one actual issue for appeal. Applicant is required to submit an appeal brief in accordance with 37 CFR 41.37. The time period for filing an appeal brief will be reset to be one month from mailing this decision, or the balance of the two-month time period running from the receipt of the notice of appeal, whichever is greater. Further, the time period for filing of the appeal brief is extendible under 37 CFR 1.136 based upon the mail date of this decision or the receipt date of the notice of appeal, as applicable.

- The panel has determined the status of the claim(s) is as follows:  
 Claim(s) allowed: \_\_\_\_\_.  
 Claim(s) objected to: \_\_\_\_\_.  
 Claim(s) rejected: \_\_\_\_\_.  
 Claim(s) withdrawn from consideration: \_\_\_\_\_.

3.  **Allowable application** – A conference has been held. The rejection is withdrawn and a Notice of Allowance will be mailed. Prosecution on the merits remains closed. No further action is required by applicant at this time.

4.  **Reopen Prosecution** – A conference has been held. The rejection is withdrawn and a new Office action will be mailed. No further action is required by applicant at this time.

All participants:

- (1) WILLIAM TROST IV. (3)\_\_\_\_\_.
- (2) Siming Liu. (4)\_\_\_\_\_.

/William Trost/  
 Supervisory Patent Examiner, Art  
 Unit 2472

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No.: 11/354,569	Confirmation No.: 6912
Applicant: Ping Pan	Group Art Unit: 2472
Filing Date: February 14, 2006	Examiner: Liu, Siming
Docket No.: 002.P045	
Customer No.: 65638	<b>Pre-Appeal Brief Request for Review TO FINAL OFFICE ACTION MAILED JULY 30, 2010</b>
For: Pseudowire Protection Using a Standby Pseudowire	<b>SUBMITTED THROUGH EFS-WEB</b>

**PRE-APPEAL BRIEF REQUEST FOR REVIEW**

Dear Panel:

In response to the Final Office Action mailed July 30, 2010, Applicant respectfully submits this Pre-Appeal Brief Request for Review and asks that the Panel consider the following remarks.

**REMARKS**

The above-referenced patent application has been reviewed in light of the Final Office Action mailed **July 30, 2010** (the “ Final Action”) and an Advisory Action mailed October 14, 2010 (the “Advisory Action”). In the Final Action, claims 1-5, 10-13, 17, 18, 23, 25 and 27 were rejected under 35 U.S.C. § 103(a) as being unpatentable over a publication entitled “The LSP Protection/Restoration Mechanism in GMPLS” by Chen (“Chen”) in view of US 2006/0047851 to Voit *et al.* (“Voit”), further in view of US 2004/0133692 to Blanchet *et al.* (“Blanchet”). Also, claims 22, 24 and 26 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Chen, in view of Voit and Blanchet and further in view of US 2006/0046658 to Cruz *et al.* (“Cruz”). The Advisory Action maintained all of the above-mentioned rejections.

**Current Status of Claims:**

Claims 1-5, 10-13, 17, 18 and 22-27 remain pending.

**Rejection of claims 1-5, 10-13, 17, 18, 23, 25 and 27 under 35 U.S.C. § 103(a):**

A portion of claim 1, as previously presented, recites:

“sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;...

*determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.”*

Emphasis added.

Applicants respectfully submit that Chen fails to describe at least the above-emphasized portions of claim 1. Chen describes the use of label distribution protocols to indicate link protection types during LSP signaling. (See page 21, paragraph 4). Chen also mentions that an LSP may have two roles: primary or secondary (backup). (See page 21, paragraph 5). However, Chen states that “[t]he resources allocated for a backup LSP are **not used until the primary LSP fails**.” (See page 21, paragraph 5, emphasis added). Further, Chen describes a type of link protection in which **backup links will not transport traffic** and that resources allocated for the

backup links can be used by **other LSPs** that have lower priorities. (See page 56, 2<sup>nd</sup> full paragraph). Chen also mentions that traffic is switched over from the primary link to the backup link when the primary link fails and graphically depicts the switch over in Figure 4.6 – see below. (See page 56, 2<sup>nd</sup> full paragraph).

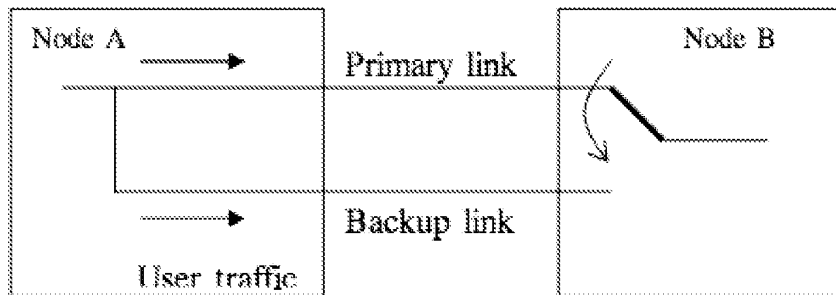


Figure 4.6: Dedicated 1+1 link protection

Further, contrary to what is stated in the Final Action, Chen describes **resource allocations** as **having priorities** and not the backup LSP. (See page 21, paragraph 5) Also, as mentioned above, Chen discloses that the resources allocated to the backup LSP may be used by other LSPs that have lower priorities until the primary fails. As stated in Chen, “[a]t that time, all the [other] LSPs using the resource allocated for the backup LSP must be preempted.” (See page 21, paragraph 5) Therefore, Chen merely describes **preempting the use** of these **prioritized resources** by the other LSPs. Since Chen describes the backup LSP as not transporting traffic until the primary LSP fails and merely describes preempting the use of prioritized resources, Chen does not describe or even suggest “determining whether to **preempt existing traffic** on the standby Pseudowire, wherein the determination is based, at least in part, **on the priority for the standby Pseudowire.**” (Emphasis added).

Voit and Blanchet were both cited in the Action to address admitted deficiencies in Chen. However, neither Voit nor Blanchet were cited as describing the above-emphasized portions of claim 1. Applicant submits that for at least the above-emphasized portions of claim 1, the Examiner has failed to show that Chen in view of Voit and further in view of Blanchet supports a *prima facie* 35 U.S.C. §103(a) rejection of claim 1. Therefore, Applicant requests that the rejection of claim 1 be withdrawn.

Independent claims 11 and 17 include similar elements to those mentioned above for claim 1. Additionally, claims 2-4, 7, 12, 15, 18, 23, 25 and 27 depend from one of claims 1, 11 or 17. Thus, Applicant requests that the 35 U.S.C. §103(a) rejections of 2-4, 7, 11, 12, 15, 17, 18, 23, 25 and 27 also be withdrawn.

**Rejection of claims 22, 24 and 26 under 35 U.S.C. § 103(a):**

Claims 22, 24 and 26 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Chen, in view of Voit and Blanchet and further in view of Cruz. Claims 1, 11 and 17 are base claims for claim 22, 24 and 26, respectively. As a result, for the same reasons mentioned above for claim 1, the Examiner has failed to show that Chen, Voit, and Blanchet support a *prima facie* 35 U.S.C. §103(a) rejection of claims 22, 24 and 26. Also, Cruz does not cure the above-stated deficiencies of Chen, Voit and Blanchet. Thus, Applicant requests that the 35 U.S.C. §103(a) rejections of claims 22, 24 and 26 be withdrawn.

**Conclusion:**

Applicant respectfully submits that the Examiner has failed to support *prima facie* rejections under § 103(a) and thus allowance of all pending claims is requested.

Respectfully submitted,

Date: October 21, 2010

by: /Ted A. Crawford/Reg. No. 50,610/  
Ted A. Crawford  
Reg. No. 50,610

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

<b>PRE-APPEAL BRIEF REQUEST FOR REVIEW</b>		Docket Number (Optional) 002.P045
I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to "Mail Stop AF, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450" [37 CFR 1.8(a)] on <u>N/A Submitted via EFS-Web</u>  Signature _____  Typed or printed name _____	Application Number 11/354,569  First Named Inventor Ping Pan  Art Unit 2472	Filed 2/14/2006  Examiner Liu, Siming
Applicant requests review of the final rejection in the above-identified application. No amendments are being filed with this request.  This request is being filed with a notice of appeal.  The review is requested for the reason(s) stated on the attached sheet(s). Note: No more than five (5) pages may be provided.		
I am the  <input type="checkbox"/> applicant/inventor.  <input type="checkbox"/> assignee of record of the entire interest. See 37 CFR 3.71. Statement under 37 CFR 3.73(b) is enclosed. (Form PTO/SB/96)  <input checked="" type="checkbox"/> attorney or agent of record. Registration number <u>50,610</u>  <input type="checkbox"/> attorney or agent acting under 37 CFR 1.34. Registration number if acting under 37 CFR 1.34 _____	_____ /Ted A. Crawford/ Signature _____ Ted A. Crawford Typed or printed name _____ 503-551-9442 Telephone number _____ 10/21/2010 Date	
NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below*.		
<input type="checkbox"/> *Total of _____ forms are submitted.		

This collection of information is required by 35 U.S.C. 132. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11, 1.14 and 41.6. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop AF, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

## Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.



<b>NOTICE OF APPEAL FROM THE EXAMINER TO THE BOARD OF PATENT APPEALS AND INTERFERENCES</b>		Docket Number (Optional) 002.P045	
I hereby certify that this correspondence is being facsimile transmitted to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to "Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450" [37 CFR 1.8(a)] N/A Submitted via EFS-Web on _____  Signature _____  Typed or printed name _____		In re Application of <b>Pseudowire Protection Using a Standby Pseudowire</b>	
		Application Number 11/354,569	Filed 2/14/2006
		For <b>Ping Pan</b>	
		Art Unit 2472	Examiner Liu, Siming
Applicant hereby <b>appeals</b> to the Board of Patent Appeals and Interferences from the last decision of the examiner.			
The fee for this Notice of Appeal is (37 CFR 41.20(b)(1))		\$ 540.00 _____	
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. Therefore, the fee shown above is reduced by half, and the resulting fee is:		\$ _____	
<input type="checkbox"/> A check in the amount of the fee is enclosed.			
<input checked="" type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.			
<input type="checkbox"/> The Director has already been authorized to charge fees in this application to a Deposit Account.			
<input type="checkbox"/> The Director is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. _____.			
<input type="checkbox"/> A petition for an extension of time under 37 CFR 1.136(a) (PTO/SB/22) is enclosed.			
<b>WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.</b>			
I am the			
<input type="checkbox"/> applicant/inventor.		/Ted A. Crawford/ _____ Signature	
<input type="checkbox"/> assignee of record of the entire interest. See 37 CFR 3.71. Statement under 37 CFR 3.73(b) is enclosed. (Form PTO/SB/96)		Ted A. Crawford _____ Typed or printed name	
<input checked="" type="checkbox"/> attorney or agent of record. Registration number 50,610 _____.		503-551-9442 _____ Telephone number	
<input type="checkbox"/> attorney or agent acting under 37 CFR 1.34. Registration number if acting under 37 CFR 1.34. _____		10/21/2010 _____ Date	
NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below*.			

<input type="checkbox"/> *Total of _____ forms are submitted.
---

This collection of information is required by 37 CFR 41.31. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11, 1.14 and 41.6. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

## Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Electronic Patent Application Fee Transmittal				
<b>Application Number:</b>	11354569			
<b>Filing Date:</b>	14-Feb-2006			
<b>Title of Invention:</b>	Pseudowire protection using a standby pseudowire			
<b>First Named Inventor/Applicant Name:</b>	Ping Pan			
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt			
<b>Attorney Docket Number:</b>	002.P045			
Filed as Large Entity				
<b>Utility under 35 USC 111(a) Filing Fees</b>				
Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
<b>Basic Filing:</b>				
<b>Pages:</b>				
<b>Claims:</b>				
<b>Miscellaneous-Filing:</b>				
<b>Petition:</b>				
<b>Patent-Appeals-and-Interference:</b>				
Notice of appeal	1401	1	540	540
<b>Post-Allowance-and-Post-Issuance:</b>				
<b>Extension-of-Time:</b>				

Description	Fee Code	Quantity	Amount	Sub-Total in USD(\$)
<b>Miscellaneous:</b>				
<b>Total in USD (\$)</b>				<b>540</b>

## Electronic Acknowledgement Receipt

<b>EFS ID:</b>	8678018
<b>Application Number:</b>	11354569
<b>International Application Number:</b>	
<b>Confirmation Number:</b>	6912
<b>Title of Invention:</b>	Pseudowire protection using a standby pseudowire
<b>First Named Inventor/Applicant Name:</b>	Ping Pan
<b>Customer Number:</b>	65638
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt
<b>Filer Authorized By:</b>	Ted A. Crawford
<b>Attorney Docket Number:</b>	002.P045
<b>Receipt Date:</b>	21-OCT-2010
<b>Filing Date:</b>	14-FEB-2006
<b>Time Stamp:</b>	18:55:09
<b>Application Type:</b>	Utility under 35 USC 111(a)

### Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	\$540
RAM confirmation Number	6776
Deposit Account	
Authorized User	

### File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip (if appl.)	Pages (if appl.)
-----------------	----------------------	-----------	-------------------------------------	-----------------------------	------------------

JUNIPER Exhibit 1003

App. 3, pg. 91

'652 File History 091

1	Pre-Brief Conference request	Pre_Appeal_Brief_Request_for_Review_002_P045_Remarks.pdf	137455 cd1fde7b8753c04973e129350e20583040c a9de5	no	4
<b>Warnings:</b>					
<b>Information:</b>					
2	Pre-Brief Conference request	Pre_Appeal_Brief_Request_for_Review_002_P045.pdf	234007 94fbc9ec3e61ef7ea7239f5f04655f86179b b19	no	2
<b>Warnings:</b>					
<b>Information:</b>					
3	Notice of Appeal Filed	Notice_of_Appeal_11_354569.pdf	244273 856e6391cb769aec8eb70c06074428d9ec6 47576	no	2
<b>Warnings:</b>					
<b>Information:</b>					
4	Fee Worksheet (PTO-875)	fee-info.pdf	29410 668d6659849b4f4fe03e24915455a11e3c54 d64ac	no	2
<b>Warnings:</b>					
<b>Information:</b>					
<b>Total Files Size (in bytes):</b>				645145	
<p><b>This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.</b></p> <p><b><u>New Applications Under 35 U.S.C. 111</u></b>  <b>If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.</b></p> <p><b><u>National Stage of an International Application under 35 U.S.C. 371</u></b>  <b>If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.</b></p> <p><b><u>New International Application Filed with the USPTO as a Receiving Office</u></b>  <b>If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.</b></p>					



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	002.P045	6912
65638	7590	10/14/2010	EXAMINER	
OMIKRON IP LAW GROUP 16325 Boones Ferry Rd. SUITE 204 LAKE OSWEGO, OR 97035			LIU, SIMING	
			ART UNIT	PAPER NUMBER
			2472	
			MAIL DATE	DELIVERY MODE
			10/14/2010	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Advisory Action Before the Filing of an Appeal Brief</b>	<b>Application No.</b> 11/354,569	<b>Applicant(s)</b> PAN, PING	
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 2472	

**--The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**

THE REPLY FILED 27 September 2010 FAILS TO PLACE THIS APPLICATION IN CONDITION FOR ALLOWANCE.

1.  The reply was filed after a final rejection, but prior to or on the same day as filing a Notice of Appeal. To avoid abandonment of this application, applicant must timely file one of the following replies: (1) an amendment, affidavit, or other evidence, which places the application in condition for allowance; (2) a Notice of Appeal (with appeal fee) in compliance with 37 CFR 41.31; or (3) a Request for Continued Examination (RCE) in compliance with 37 CFR 1.114. The reply must be filed within one of the following time periods:

- a)  The period for reply expires \_\_\_\_\_ months from the mailing date of the final rejection.
- b)  The period for reply expires on: (1) the mailing date of this Advisory Action, or (2) the date set forth in the final rejection, whichever is later. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of the final rejection.
- Examiner Note: If box 1 is checked, check either box (a) or (b). ONLY CHECK BOX (b) WHEN THE FIRST REPLY WAS FILED WITHIN TWO MONTHS OF THE FINAL REJECTION. See MPEP 706.07(f).

Extensions of time may be obtained under 37 CFR 1.136(a). The date on which the petition under 37 CFR 1.136(a) and the appropriate extension fee have been filed is the date for purposes of determining the period of extension and the corresponding amount of the fee. The appropriate extension fee under 37 CFR 1.17(a) is calculated from: (1) the expiration date of the shortened statutory period for reply originally set in the final Office action; or (2) as set forth in (b) above, if checked. Any reply received by the Office later than three months after the mailing date of the final rejection, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**NOTICE OF APPEAL**

2.  The Notice of Appeal was filed on \_\_\_\_\_. A brief in compliance with 37 CFR 41.37 must be filed within two months of the date of filing the Notice of Appeal (37 CFR 41.37(a)), or any extension thereof (37 CFR 41.37(e)), to avoid dismissal of the appeal. Since a Notice of Appeal has been filed, any reply must be filed within the time period set forth in 37 CFR 41.37(a).

**AMENDMENTS**

3.  The proposed amendment(s) filed after a final rejection, but prior to the date of filing a brief, will not be entered because
- (a)  They raise new issues that would require further consideration and/or search (see NOTE below);
- (b)  They raise the issue of new matter (see NOTE below);
- (c)  They are not deemed to place the application in better form for appeal by materially reducing or simplifying the issues for appeal; and/or
- (d)  They present additional claims without canceling a corresponding number of finally rejected claims.

NOTE: \_\_\_\_\_. (See 37 CFR 1.116 and 41.33(a)).

4.  The amendments are not in compliance with 37 CFR 1.121. See attached Notice of Non-Compliant Amendment (PTOL-324).
5.  Applicant's reply has overcome the following rejection(s): \_\_\_\_\_.
6.  Newly proposed or amended claim(s) \_\_\_\_\_ would be allowable if submitted in a separate, timely filed amendment canceling the non-allowable claim(s).
7.  For purposes of appeal, the proposed amendment(s): a)  will not be entered, or b)  will be entered and an explanation of how the new or amended claims would be rejected is provided below or appended.
- The status of the claim(s) is (or will be) as follows:  
 Claim(s) allowed: \_\_\_\_\_.  
 Claim(s) objected to: \_\_\_\_\_.  
 Claim(s) rejected: 1-5, 10-13, 17, 18 and 22-27.  
 Claim(s) withdrawn from consideration: 6-9, 14-16 and 19-21.

**AFFIDAVIT OR OTHER EVIDENCE**

8.  The affidavit or other evidence filed after a final action, but before or on the date of filing a Notice of Appeal will not be entered because applicant failed to provide a showing of good and sufficient reasons why the affidavit or other evidence is necessary and was not earlier presented. See 37 CFR 1.116(e).
9.  The affidavit or other evidence filed after the date of filing a Notice of Appeal, but prior to the date of filing a brief, will not be entered because the affidavit or other evidence failed to overcome all rejections under appeal and/or appellant fails to provide a showing of good and sufficient reasons why it is necessary and was not earlier presented. See 37 CFR 41.33(d)(1).
10.  The affidavit or other evidence is entered. An explanation of the status of the claims after entry is below or attached.

**REQUEST FOR RECONSIDERATION/OTHER**

11.  The request for reconsideration has been considered but does NOT place the application in condition for allowance because:  
See Continuation Sheet.
12.  Note the attached Information *Disclosure Statement*(s). (PTO/SB/08) Paper No(s). \_\_\_\_\_
13.  Other: \_\_\_\_\_.

/William Trost/  
Supervisory Patent Examiner, Art Unit 2472

/S. L./  
Examiner, Art Unit 2472



Continuation of 11. does NOT place the application in condition for allowance because: Regarding applicant's argument that the prior art does not disclose the limitation "determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire". Examiner respectfully disagrees. Applicant pointed out that the resource is idle, no traffic can exist on the backup LSP. According to page 21, last paragraph of Chen, Chen discloses that "resource allocated for a backup LSP may be used by an LSP that has lower priority until primary LSP fails". The quote from Chen indicates that the backup path is not idle, it can be utilized by another LSP until primary LSP fails. Chen also disclosed the switch over step when the primary LSP fails. The switch over is also based on the priority. The LSP utilize the backup path has a lower priority. When the primary fails, the traffic is switched over to the backup. "At that time, all the LSPs using the resource allocated for the backup LSP must be preempted".

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No.: 11/354,569

Confirmation No.: 6912

Applicant: Ping Pan

Group Art Unit: 2472

Filing Date: February 14, 2006

Examiner: Liu, Siming

Docket No.: 002.P045

Customer No.: 65638

**RESPONSE  
TO FINAL OFFICE ACTION  
MAILED JULY 30, 2010**

For: Pseudowire Protection Using a Standby  
Pseudowire

**SUBMITTED THROUGH EFS-WEB**

**RESPONSE AFTER FINAL**

Dear Sir:

In response to the Final Office Action mailed July 30, 2010, Applicant respectfully requests that the Examiner favorably consider the following remarks.

Remarks begin at page 2 of this paper.

**REMARKS**

The above-referenced patent application has been reviewed in light of the Final Office Action mailed **July 30, 2010** (the “Action”). In the Action, claims 1-5, 10-13, 17, 18, 23, 25 and 27 were rejected under 35 U.S.C. § 103(a) as being unpatentable over a publication entitled “The LSP Protection/Restoration Mechanism in GMPLS” by Chen (“Chen”) in view of US 2006/0047851 to Voit *et al.* (“Voit”), further in view of US 2004/0133692 to Blanchet *et al.* (“Blanchet”). Also, claims 22, 24 and 26 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Chen, in view of Voit and Blanchet and further in view of US 2006/0046658 to Cruz *et al.* (“Cruz”).

**Current Status of Claims:**

Claims 1-5, 10-13, 17, 18 and 22-27 remain pending.

**Rejection of claims 1-5, 10-13, 17, 18, 23, 25 and 27 under 35 U.S.C. § 103(a):**

A portion of claim 1, as previously presented, recites:

“sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;...

*determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.”*

Emphasis added.

Applicants respectfully submit that Chen fails to describe at least the above-emphasized portions of claim 1. Chen describes the use of label distribution protocols to indicate link protection types during LSP signaling. (See page 21, paragraph 4). Chen also mentions that an LSP may have two roles: primary or secondary (backup). (See page 21, paragraph 5). However, Chen states that “[t]he resources allocated for a backup LSP are **not used until the primary LSP fails.**” (See page 21, paragraph 5, emphasis added). Applicants submit that if the resources allocated to the backup LSP are not used until the primary LSP fails, then the backup LSP is idle. Since the backup LSP is idle, no traffic can exist on the backup LSP.

Further, contrary to what is stated in the Action, Chen describes **resource allocations** as **having priorities** and not the backup LSP. (See page 21, paragraph 5) Also, Chen mentions that the resources allocated to the backup LSP may be used by other LSPs until the primary fails. As stated in Chen, “[a]t that time, all the [other] LSPs using the resource allocated for the backup LSP must be preempted.” (See page 21, paragraph 5) Therefore, Chen describes **preempting the use** of these **prioritized resources** by the other LSPs. Since Chen describes the backup LSP as being idle (i.e., no existing traffic) and also only describes preempting the use of prioritized resources, Chen does not describe or even suggest “determining whether to **preempt existing traffic** on the standby Pseudowire, wherein the determination is based, at least in part, **on the priority for the standby Pseudowire.**” (Emphasis added).

Voit and Blanchet were both cited in the Action to address admitted deficiencies in Chen. However, neither Voit nor Blanchet were cited as describing the above-emphasized portions of claim 1. Applicant submits that for at least the above-emphasized portions of claim 1, Chen in view of Voit and further in view of Blanchet do not support a *prima facie* 35 U.S.C. §103(a) rejection of claim 1. Therefore, Applicant requests that the rejection of claim 1 be withdrawn.

Independent claims 11 and 17 include similar elements to those mentioned above for claim 1. Additionally, claims 2-4, 7, 12, 15, 18, 23, 25 and 27 depend from one of claims 1, 11 or 17. Thus, Applicant requests that the 35 U.S.C. §103(a) rejections of 2-4, 7, 11, 12, 15, 17, 18, 23, 25 and 27 also be withdrawn.

**Rejection of claims 22, 24 and 26 under 35 U.S.C. § 103(a):**

Claims 22, 24 and 26 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Chen, in view of Voit and Blanchet and further in view of Cruz. Claims 1, 11 and 17 are base claims for claim 22, 24 and 26, respectively. As a result, for the same reasons mentioned above for claim 1, Chen, Voit, and Blanchet do not support a *prima facie* 35 U.S.C. §103(a) rejection of claims 22, 24 and 26. Also, Cruz does not cure the above-stated deficiencies of Chen, Voit and Blanchet. Thus, Applicant requests that the 35 U.S.C. §103(a) rejections of claims 22, 24 and 26 be withdrawn.

**Conclusion:**

Applicant respectfully submits that claims 1-5, 10-13, 17, 18 and 22-27 are in condition for allowance and such action is earnestly solicited. ***The Examiner is respectfully requested to contact the undersigned by telephone at (503) 551-9442 if it is believed that such contact would further the examination of the present application.***

Respectfully submitted,

Date: September 27, 2010

by: /Ted A. Crawford/Reg. No. 50,610/  
Ted A. Crawford  
Reg. No. 50,610

## Electronic Acknowledgement Receipt

<b>EFS ID:</b>	8502540
<b>Application Number:</b>	11354569
<b>International Application Number:</b>	
<b>Confirmation Number:</b>	6912
<b>Title of Invention:</b>	Pseudowire protection using a standby pseudowire
<b>First Named Inventor/Applicant Name:</b>	Ping Pan
<b>Customer Number:</b>	65638
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt
<b>Filer Authorized By:</b>	Ted A. Crawford
<b>Attorney Docket Number:</b>	002.P045
<b>Receipt Date:</b>	27-SEP-2010
<b>Filing Date:</b>	14-FEB-2006
<b>Time Stamp:</b>	13:29:10
<b>Application Type:</b>	Utility under 35 USC 111(a)

### Payment information:

Submitted with Payment	no
------------------------	----

### File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Amendment After Final	Response_to_Final_002_P045.pdf	130798 677b7f1c1f6ff44629b20f938f46629491ac5689	no	4

### Warnings:

### Information:

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

**New Applications Under 35 U.S.C. 111**

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

**National Stage of an International Application under 35 U.S.C. 371**

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

**New International Application Filed with the USPTO as a Receiving Office**

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	002.P045	6912
65638	7590	07/30/2010	EXAMINER	
OMIKRON IP LAW GROUP 16325 Boones Ferry Rd. SUITE 204 LAKE OSWEGO, OR 97035			LIU, SIMING	
			ART UNIT	PAPER NUMBER
			2472	
			MAIL DATE	DELIVERY MODE
			07/30/2010	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.



<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	11/354,569	PAN, PING	
	<b>Examiner</b>	<b>Art Unit</b>	
	SIMING LIU	2472	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1)  Responsive to communication(s) filed on 21 April 2010.
- 2a)  This action is **FINAL**.                      2b)  This action is non-final.
- 3)  Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4)  Claim(s) 1-27 is/are pending in the application.
  - 4a) Of the above claim(s)        is/are withdrawn from consideration.
- 5)  Claim(s)        is/are allowed.
- 6)  Claim(s) 1-5, 10-13, 17-18, 22-27 is/are rejected.
- 7)  Claim(s)        is/are objected to.
- 8)  Claim(s)        are subject to restriction and/or election requirement.

**Application Papers**

- 9)  The specification is objected to by the Examiner.
- 10)  The drawing(s) filed on        is/are: a)  accepted or b)  objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11)  The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12)  Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    - a)  All    b)  Some \* c)  None of:
    - 1.  Certified copies of the priority documents have been received.
    - 2.  Certified copies of the priority documents have been received in Application No.       .
    - 3.  Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1)  Notice of References Cited (PTO-892)
- 2)  Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3)  Information Disclosure Statement(s) (PTO/SB/08)  
 Paper No(s)/Mail Date       .
- 4)  Interview Summary (PTO-413)  
 Paper No(s)/Mail Date.       .
- 5)  Notice of Informal Patent Application
- 6)  Other:       .

## DETAILED ACTION

### *Response to Arguments*

1. Applicant's arguments have been considered but are moot in view of the new ground(s) of rejection. Applicant amended the independent claims significantly, which necessitates the new ground of rejection.

### *Claim Rejections - 35 USC § 103*

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

3. Claims 1-5, 10-13, 17, 23, 25, 27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chen, "The LSP Protection/Restoration Mechanism in GMPLS", in view of Voit US 2006/0047851 A1, further in view of Blanchet US 2004/0133692 A1.
4. Regarding claims 1, 11, 17, Chen teaches a method/system/computer product of providing protection to network traffic, comprising:  
sending a ... protection configuration parameter for configuring a standby ... between a source node and a destination node, the ... protection configuration parameter

indicating a protection property associated with the standby ... (page 21, paragraph 4, lines 1-2: "label distribution protocols may carry the link protection type", the link protection type is protection configuration parameter), the protection property including a priority for the standby ... (page 21, paragraph 5, lines 2-6: "the resource allocation has priorities (carried by the signaling protocol), the resources allocated for a backup LSP may be used by an LSP that ...until primary LSP fails and the traffic is switched over to the backup", it's noted the priority carried by the signaling protocol is the protection configuration parameter); ... and accepting the ... protection configuration parameter by the destination node;

using the standby ... that is configured based at least in part on the ... protection configuration parameter; and determining whether to preempt existing traffic on the standby ..., wherein the determination is based, at least in part, on the priority for the standby ... (Page 21, last paragraph: lines 3-7: "Because the resource allocation has priorities (carried by the signaling protocol), ... all the LSPs using the resource allocated for the backup LSP must be preempt", it's noted the configuration parameter is carried in the signaling protocol)

Chen teaches path protection/restoration in GMPLS, but it doesn't expressly teach Pseudowire and Pseudowire protection.

However, Voit teaches Pseudowire (Voit, page 2, [0011], lines 2-7) and Pseudowire protection (Voit, page 4, [0046], lines 1-3: "a network topology is provided with redundant pseudowire connections ...").

At the time of the invention was made, it would have been obvious to a person of ordinary skill in the art to apply the protection/restoration mechanism disclosed by Chen in Pseudowires environment. Both MPLS and Pseudowire are point-to-point virtual link. Voit also teaches providing data traffic protection for primary Pseudowire path (Voit, page 4, [0046], lines 1-3). Therefore, the combination is to apply a known technique to a similar system to improve its reliability. Both Chen and Voit are in the same field of endeavor (network transfer) and are directed to the same problem sought to be solved (data traffic protection).

Chen in view of Voit doesn't expressly teach that receiving a configuration acknowledgement indicating whether the configuration parameter has been accepted by the destination node.

Blanchet teaches that receiving a configuration acknowledgement indicating whether the configuration parameter has been accepted by the destination node (Blanchet, page 4, [0035], lines 2-4).

At the time of the invention was made, it would have been obvious to a person of ordinary skill in the art to modify the system to send an ACK indicating the acceptance of the configuration parameters in the system disclosed by Chen in view of Voit in order to makes the system more reliable. Both Chen in view of Voit and Blanchet are in the same field of endeavor (Network transfer).

5. Regarding claims 2, 12, Chen in view of Voit and Blanchet further teaches the standby Pseudowire (Voit, [0002], lines 1-2) is configured to provide protection to at

least one primary Pseudowire (Chen, page 21, last paragraph, "There are two LSP roles: primary or secondary (backup). The GMPLS signaling protocol carries a flag that indicates ... the resource allocated for a backup LSP may be used by an LSP that has lower priority until the primary LSP fails and the traffic is switched over to the backup").

6. Regarding claim 3, Chen in view of Voit and Blanchet further teaches the standby Pseudowire (Voit, [0002], lines 1-2) is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire (Chen, page 21, last paragraph, "There are two LSP roles: primary or secondary (backup). The GMPLS signaling protocol carries a flag that indicates ... the resource allocated for a backup LSP may be used by an LSP that has lower priority until the primary LSP fails and the traffic is switched over to the backup").

7. Regarding claim 4, Chen in view of Voit and Blanchet further teaches wherein the standby Pseudowire is dynamically selected from a plurality of connections (Chen, page 22, last paragraph: the backup path is dynamically chosen from a plurality of connection).

8. Regarding claims 5, 13, 18, Chen in view of Voit and Blanchet further teaches the protection property configuration parameter further includes at least one of a domain

type, a protection type or a protection scheme (Chen, page 21, 4<sup>th</sup> paragraph, lines “during LSP signaling in GMPLS, label distribution protocols may carry the link protection type”, the limitations are presented in alternative form, therefore only one of them needs to be addressed to meet the claim limitation).

9. Regarding claim 10, Chen in view of Voit and Blanchet further teaches the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP) (Chen, page 21, 4<sup>th</sup> paragraph, lines “during LSP signaling in GMPLS, label distribution protocols may carry the link protection type”).

10. Regarding claim 23, 25, 27, Chen in view of Voit and Blanchet further teaches the protection scheme indicates at least one of the following:

a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires (Chen, page 17: “dedicated 1+1”); a 1:1 protection scheme, wherein one Pseudowire is used to protect another Pseudowire (Chen, page 17, “dedicated 1+1”); a 1 :N protection scheme, wherein one Pseudowire is used to protect N other Pseudowires (Chen, page 54: “1:N protection”);

or an M:N protection scheme, wherein M Pseudowires are used to protect N other Pseudowires (Chen, page 53, “M:N protection”).

11. Claims 22, 24, 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over under 35 U.S.C. 103(a) as being unpatentable over Chen, in view of Voit and Blanchet, further in view of Cruz, US 2006/0046658 A1.

12. Regarding claims 22, 24, 26, Chen in view of Voit, Blanchet teaches all of the limitations except that domain type indicates whether the Pseudowire is configured in a single-hop environment where the Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the Pseudowire includes a plurality of nodes coupled to several carrier networks.

Cruz teaches a domain type indicates whether the Pseudowire is configured in a single-hop environment where the Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the Pseudowire includes a plurality of nodes coupled to several carrier networks. (Cruz, page 1, [0017], line 2: According to the specification of the application, domain type is about whether the network is either multi-hop or single hop).

At the time of the invention was made, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include domain type. The reason is that by including domain type in the configuration parameter, it would be more accurate to select a desire standby path, given that you have more information about the network. The method of change the configuration parameter by including the domain type of Chen in view of Voit, Blanchet was within the ordinary ability of one of ordinary skill in the art based on the teachings of Cruz.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Chen, Voit, Blanchet and Cruz to obtain the invention as specified in claims 22, 24, 26.

***Conclusion***

13. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIMING LIU whose telephone number is (571)270-3859. The examiner can normally be reached on Monday-Friday 8:30am-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, William Trost can be reached on 571-272-7872. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.



Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/S. L./  
Examiner, Art Unit 2472

/William Trost/  
Supervisory Patent Examiner, Art  
Unit 2472

<b>Notice of References Cited</b>	Application/Control No. 11/354,569	Applicant(s)/Patent Under Reexamination PAN, PING	
	Examiner SIMING LIU	Art Unit 2472	Page 1 of 1

**U.S. PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A US-			
	B US-			
	C US-			
	D US-			
	E US-			
	F US-			
	G US-			
	H US-			
	I US-			
	J US-			
	K US-			
	L US-			
	M US-			

**FOREIGN PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N				
	O				
	P				
	Q				
	R				
	S				
	T				

**NON-PATENT DOCUMENTS**


*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)				
*	U			Ziying Chen: "The LSP Protection/Restoration Mechanism in GMPLS" Internet Citatio (Online) October 2002 (2002-10-01), XP002239552 Retrieved from the ineternet URL: <a href="http://www.site.uottawa.ca/~bochmann/dsrg/PublicDocuments/Master-theses/Chen,%20Ziying%20%20-%20202002.pdf">http://www.site.uottawa.ca/~bochmann/dsrg/PublicDocuments/Master-theses/Chen,%20Ziying%20%20-%20202002.pdf</a>	
	V				
	W				
	X				

\*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)  
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

<b>Index of Claims</b>  	<b>Application/Control No.</b> 11354569	<b>Applicant(s)/Patent Under Reexamination</b> PAN, PING
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 2472

✓	<b>Rejected</b>	-	<b>Cancelled</b>	N	<b>Non-Elected</b>	A	<b>Appeal</b>
=	<b>Allowed</b>	÷	<b>Restricted</b>	I	<b>Interference</b>	O	<b>Objected</b>

<input type="checkbox"/> Claims renumbered in the same order as presented by applicant		<input type="checkbox"/> CPA		<input type="checkbox"/> T.D.		<input type="checkbox"/> R.1.47			
CLAIM		DATE							
Final	Original	10/30/2008	06/17/2009	01/14/2010	07/16/2010				
	1	✓	✓	✓	✓				
	2	✓	✓	✓	✓				
	3	✓	✓	✓	✓				
	4	✓	✓	✓	✓				
	5	✓	✓	✓	✓				
	6	✓	✓	✓	-				
	7	✓	✓	✓	-				
	8	✓	✓	✓	-				
	9	✓	✓	✓	-				
	10	✓	✓	✓	✓				
	11	✓	✓	✓	✓				
	12	✓	✓	✓	✓				
	13	✓	✓	✓	✓				
	14	✓	✓	✓	-				
	15	✓	✓	✓	-				
	16	✓	✓	✓	-				
	17	✓	✓	✓	✓				
	18	✓	✓	✓	✓				
	19	✓	✓	✓	-				
	20	✓	✓	✓	-				
	21	✓	✓	✓	-				
	22				✓				
	23				✓				
	24				✓				
	25				✓				
	26				✓				
	27				✓				

<b>Search Notes</b>  	<b>Application/Control No.</b>  11354569	<b>Applicant(s)/Patent Under Reexamination</b>  PAN, PING
	<b>Examiner</b>  SIMING LIU	<b>Art Unit</b>  2472

SEARCHED			
Class	Subclass	Date	Examiner
370	216, 225, 228	10/30/2008	/SL/
709	220	10/30/2008	/SL/
above	update search	6/17/2009	/SL/
update search	ABOVE	1/14/2010	/SL/
update search	ABOVE	7/16/2010	/SL/

SEARCH NOTES		
Search Notes	Date	Examiner
East Class search	11/10/2008 update 6/17/2009	/SL/
Palm inventor name search	10/30/2008 update 6/17/2009	/SL/
Consulted 101 issues with Peng, John	11/10/2008	/SL/
update search: ABOVE	1/14/2010	/SL/
update search: ABOVE	7/16/2010	/SL/

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner

--	--

## EAST Search History

## EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S5	115	pseudowire	US-PGPUB; USPAT	OR	ON	2008/10/08 12:53
S6	0	pseudowire and tele	US-PGPUB; USPAT	OR	ON	2008/10/08 12:53
S7	217	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2008/10/08 13:07
S8	9	S7 with protection	US-PGPUB; USPAT	OR	ON	2008/10/08 13:08
S9	4	S7 with protection and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 13:09
S10	1	"20040223498".pn.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:07
S11	0	(pseudowire or pseudo-wire) and initiliz\$5	US-PGPUB; USPAT	OR	ON	2008/10/08 14:14
S12	133	(pseudowire or pseudo-wire) and initi\$5	US-PGPUB; USPAT	OR	ON	2008/10/08 14:15
S13	51	(pseudowire or pseudo-wire) and initi\$5 and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:15
S14	2193	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:17
S15	6	(370/216,225,228).ccls. and S7	US-PGPUB; USPAT	OR	ON	2008/10/08 14:23
S16	3	(709/220).ccls. and S7	US-PGPUB; USPAT	OR	ON	2008/10/08 14:26
S17	31	((PING) near2 (PAN)).INV.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:32
S18	2	((PING) near2 (PAN)).INV. and pseudowire	US-PGPUB; USPAT	OR	ON	2008/10/08 14:33
S19	2	((PING) near2 (PAN)).INV. and (pseudowire).clm.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:33
S20	66	S7 and (primary)	US-PGPUB; USPAT	OR	ON	2008/10/08 14:38
S21	23	S7 and (primary) and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:39
S22	75	S7 and (config\$7) and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:44
S23	2	TDM pseudowire	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:09
S24	106	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol))	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:16

S26	43	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:17
S27	11	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (standby or backup)	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:19
S28	9	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (primary or main) and (secondly or backup or standby)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 09:00
S29	43	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:34
S30	14	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:35
S31	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:38
S32	18	(pseudowire or (pseudo wire) or pseudo-wire) and ((config\$7) same (destination near5 node))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:38
S33	1	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with acknowledgement)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:41
S34	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:44
S35	8	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:44
S36	370	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:47
S37	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S33) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:48

S38	218	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2008/10/09 10:49
S39	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S38) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:49
S40	275	pseudowire or pseudo-wire or (pseudo wire)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:49
S41	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S40) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:50
S42	233	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:50
S43	27	S40 and initialization	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:54
S44	3434011	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2008/10/29 09:26
S46	1334019	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:27
S47	3640734	(alter\$7 or backup or standby)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:28
S48	29788561	@ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:28
S49	6386553	(pick\$5 or select\$5 or choos\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:30
S50	409	(S44 near7 S46) with (S49 near7 S47 near7 S44) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:31
S51	238	(S44 near7 S46) with (S49 near7 S47 near7 S44) and S48 and (priority or bandwidth)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:38
S52	25	(S44 near7 S46) with (S49 near7 S47 near7 S44) same (priority or bandwidth or parameter) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:43

S53	2289	S47 with config\$7 with (primary near7 S47)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:49
S54	159	(S47 near5 S44) with config \$7 with (primary near7 S47)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:54
S55	175	(S47 near5 S44) with config \$7 with (primary near7 S44)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:54
S56	111	(S47 near5 S44) with config \$7 with (primary near7 S44) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:56
S57	7	09/859166	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 10:26
S58	33	(restoration scheme) and ("1:N")	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 10:50
S59	4	(restoration scheme) and (priority) and (standby mode)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 12:42
S60	18	(restoration scheme) and (priority) and (config\$7 near5 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:27
S61	3706723	(send\$7 or transmit\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:30
S62	0	(source node) with S61 with (config\$7 near3 parameter \$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:31
S63	5	(source) with S61 with (config\$7 near3 parameter \$1) with (destin\$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:32
S64	569	(source) with S61 with (parameter\$1) with (destin \$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:33
S65	54	(source node) with S61 with (parameter\$1) with (destin \$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:33



S66	2959	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:53
S67	0	(ack or acknowledgement) same (config\$7 parameter \$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:53
S68	20807	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:54
S69	52906	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:54
S70	29	(ack or acknowledgement) and (restoration scheme) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:55
S71	137	S61 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:19
S72	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:29
S73	10549	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:29
S74	759	handshaking and @ad<"20050214" and (S44 with S46)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:30
S75	2	"6553034".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:32
S76	108	(virtual path) and ((protection or restoration) near5 scheme) and priority	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:34
S77	103	(virtual path) and ((protection or restoration) near5 scheme) and priority and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:34
S78	3479	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:51

S79	2628	((protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:52
S80	6	((protection or restoration) near5 parameter) with (S61) and (destin\$7 near3 node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:53
S81	26	((protection or restoration) near5 parameter) and (handshaking) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:55
S82	73	((protection or restoration) near5 parameter) and (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:04
S83	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:04
S84	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S85	15	((protection or restoration or config\$7) near5 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S86	4	((protection or restoration or config\$7) near2 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S87	651	handshaking and @ad<"20050214" and (config\$7 parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:10
S88	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:27
S89	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:27
S90	276	receiv\$7 acknowledgement destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:28
S91	0	receiv\$7 acknowledgement (parameter accept\$5 destination node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:40

S92	5	receiv\$7 acknowledgement accept\$3 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:45
S93	7	receiv\$7 acknowledgement parameter accept\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:48
S94	2	"20030117950".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:20
S95	771	(domain type) with (parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:25
S96	3	(parameter) near5 includ\$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:26
S97	3	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:27
S98	64	(parameter\$1) with (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:27
S99	0	(domain type) with (single- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:28
S100	0	(domain type) with (single near5 hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:29
S101	64	(domain type) with parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:29
S102	179	(single-hop) same (multi- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:32
S103	9	(single-hop) same (multi- hop) same (parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:33
S104	0	field with indica\$7 with ((single-hop) same (multi- hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:34

S105	147	field with indica\$7 with (topology)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:34
S106	6	field with indica\$7 with (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:36
S107	179	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:37
S108	10	(field or parameter) same ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:37
S109	283	((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:40
S111	21	(parameter or field) same S109	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:41
S112	0	S109 same (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:44
S113	134	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:46
S114	0	parameter with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:48
S115	0	(field or parameter) with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:49
S116	0	(field or parameter) with (show\$3 or indicat\$5) same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:49
S117	0	(field or parameter) with (domain type) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50

S118	68	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S119	14	(field or parameter) with (indicat\$5 or show\$5) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S120	1	(protection type) and (standby path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:44
S121	2636	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:46
S122	283	(hot and cold) same standby and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:47
S123	51	(hot and cold) and (parameter with standby) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:48
S124	20	(field with indicat\$5 with (standby mode)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:49
S126	696	config\$9 with (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:50
S127	406	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:51
S128	324	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:51
S129	194	config\$9 near3 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:52
S130	7	config\$9 near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:54
S131	9	type with (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:01

S132	4	type near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:01
S133	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:38
S135	134	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:39
S136	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:43
S137	126	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:43
S138	11	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) same (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:44
S139	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 netowrk)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:53
S140	136	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 network)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:53
S141	0	(paramete or field or bit) with indicat\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:55
S142	2	(paramete or field or bit) with indicat\$7 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:55
S143	1	(protection near3 properties) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2009/05/02 14:27
S144	7	(protection near3 (parameter or propert\$5)) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2009/05/02 14:28
S145	61963	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2009/05/02 15:12
S146	15	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2009/05/02 15:13

S147	25	(pseudowire or pseudo-wire) and (standby)	US-PGPUB; USPAT	OR	ON	2009/05/02 15:20
S148	3	(pseudowire or pseudo-wire) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 15:20
S149	12	(protection scheme) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:04
S150	1	(protection scheme) with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:06
S151	1	(protection (type or property)) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:22
S152	10	((protection (type or property)) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:23
S153	19	(( (type or property) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:26
S154	1	(protection scheme) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:27
S155	2	(protection scheme) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:27
S156	3	(protection (scheme or propert\$3 or parameter or type)) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:29
S157	20	(backup path) with (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:22
S158	0	(backup path) with (protection near3 parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:35
S159	121	(backup path) with (protection )	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:35
S160	1	(genera\$5 or configur\$5) with (backup path) with (parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:36
S161	144	(genera\$5 or configur\$5) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:37
S162	22	(genera\$5 or configur\$5) with (backup path) and (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:38
S163	0	(genera\$5 or configur\$5) with (backup path) with (base or according)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:42

S164	0	(genera\$5 or configur\$5) with (backup path) with ("base" or "according")	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:42
S165	159	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:43
S166	15	(genera\$5 or configur\$5 or setup) with (backup path) not S161	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:43
S168	2421	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S169	3849	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S170	291	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S171	5	(709/220).ccls. and S170	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S172	1	"20050226215"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2009/06/17 16:27
S173	2	"20060045028"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2009/06/17 16:30
S174	397	(pseudowire or pseudo-wire or pseudo wire)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 16:41
S175	13	(pseudowire or pseudo-wire or pseudo wire) and (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 16:41
S176	199	pseudowire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S177	366	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S178	11	S177 with protection	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S179	4	S177 with protection and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S180	1	"20040223498".pn.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S181	0	(pseudowire or pseudo-wire) and iniliz\$5	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S182	209	(pseudowire or pseudo-wire) and initi\$5	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S183	71	(pseudowire or pseudo-wire) and initi\$5 and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S184	2654	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S185	14	(370/216,225,228).ccls. and S177	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55



S186	6	(709/220).ccls. and S177	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S187	35	((PING) near2 (PAN)).INV.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S188	2	((PING) near2 (PAN)).INV. and pseudowire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S189	2	((PING) near2 (PAN)).INV. and (pseudowire).clm.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S190	100	S177 and (primary)	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S191	30	S177 and (primary) and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S192	106	S177 and (config\$7) and @ad<"20050214"	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S193	4	TDM pseudowire	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S194	189	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S195	62	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S196	14	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (standby or backup)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S197	12	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (primary or main) and (secondly or backup or standby)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S198	68	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S199	22	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S200	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55

S201	32	(pseudowire or (pseudo wire) or pseudo-wire) and ((config\$7) same (destination near5 node))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S202	4	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with acknowledgement)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S203	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S204	12	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S205	441	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S206	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S202) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S207	366	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S208	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S207) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S209	462	pseudowire or pseudo-wire or (pseudo wire)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S210	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S209) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S211	246	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S212	41	S209 and initialization	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
S213	3810399	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2010/01/15 11:55
S214	1494211	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

S215	4045518	(alter\$7 or backup or standby)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S216	29886744	@ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S217	7016429	(pick\$5 or select\$5 or choos\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S218	420	(S213 near7 S214) with (S217 near7 S215 near7 S213) and S216	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S219	246	(S213 near7 S214) with (S217 near7 S215 near7 S213) and S216 and (priority or bandwidth)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S220	28	(S213 near7 S214) with (S217 near7 S215 near7 S213) same (priority or bandwidth or parameter) and S216	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S221	2735	S215 with config\$7 with (primary near7 S215)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S222	188	(S215 near5 S213) with config\$7 with (primary near7 S215)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S223	207	(S215 near5 S213) with config\$7 with (primary near7 S213)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S224	119	(S215 near5 S213) with config\$7 with (primary near7 S213) and S216	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S225	7	09/859166	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S226	39	(restoration scheme) and ("1:N")	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S227	5	(restoration scheme) and (priority) and (standby mode)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

S228	22	(restoration scheme) and (priority) and (config\$7 near5 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S229	4138472	(send\$7 or transmit\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S230	0	(source node) with S229 with (config\$7 near3 parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S231	7	(source) with S229 with (config\$7 near3 parameter\$1) with (destin\$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S232	675	(source) with S229 with (parameter\$1) with (destin\$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S233	62	(source node) with S229 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S234	3563	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S235	0	(ack or acknowledgement) same (config\$7 parameter\$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S236	25761	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S237	55233	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S238	31	(ack or acknowledgement) and (restoration scheme) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S239	153	S229 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S240	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

S241	10895	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S242	797	handshaking and @ad<"20050214" and (S213 with S214)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S243	2	"6553034".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S244	114	(virtual path) and ((protection or restoration) near5 scheme) and priority	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S245	106	(virtual path) and ((protection or restoration) near5 scheme) and priority and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S246	4061	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S247	2702	(protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S248	7	((protection or restoration) near5 parameter) with (S229) and (destin\$7 near3 node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S249	27	((protection or restoration) near5 parameter) and (handshaking) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S250	76	((protection or restoration) near5 parameter) and (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S251	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S252	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S253	16	((protection or restoration or config\$7) near5 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

S254	4	((protection or restoration or config\$7) near2 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S255	693	handshaking and @ad<"20050214" and (config\$7 parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S256	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S257	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S258	325	receiv\$7 acknowledgement destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S259	0	receiv\$7 acknowledgement (parameter accept\$5 destination node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S260	5	receiv\$7 acknowledgement accept\$3 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S261	8	receiv\$7 acknowledgement parameter accept\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S262	2	"20030117950".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S263	999	(domain type) with (parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S264	4	(parameter) near5 includ\$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S265	4	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
S266	71	(parameter\$1) with (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55

S267	71	(domain type) with parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S268	258	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S269	11	(single-hop) same (multi-hop) same (parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S270	0	field with indica\$7 with ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S271	197	field with indica\$7 with (topology)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S272	10	field with indica\$7 with (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S273	258	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S274	18	(field or parameter) same ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S275	415	((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S276	35	(parameter or field) same S275	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S277	150	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S278	0	(field or parameter) with (domain type) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S279	71	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

S280	15	(field or parameter) with (indicat\$5 or show\$5) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S281	1	(protection type) and (standby path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S282	2982	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S283	292	(hot and cold) same standby and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S284	52	(hot and cold) and (parameter with standby) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S285	21	(field with indicat\$5 with (standby mode)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S286	725	config\$9 with (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S287	421	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S288	336	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S289	203	config\$9 near3 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S290	7	config\$9 near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S291	9	type with (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S292	4	type near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55



S293	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S294	150	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S295	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S296	142	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S297	14	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) same (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S298	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 netowrk)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S299	193	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 network)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S300	0	(paramete or field or bit) with indicat\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S301	13	(paramete or field or bit) with indicat\$7 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
S302	1	(protection near3 properties) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2010/01/15 11:55
S303	8	(protection near3 (parameter or propert\$5)) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2010/01/15 11:55
S304	70110	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/01/15 11:55
S305	16	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/01/15 11:55
S306	37	(pseudowire or pseudo-wire) and (standby)	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55

S307	5	(pseudowire or pseudo-wire) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S308	12	(protection scheme) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S309	1	(protection scheme) with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S310	1	(protection (type or property)) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S311	11	((protection (type or property)) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S312	20	(( (type or property) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S313	1	(protection scheme) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S314	2	(protection scheme) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S315	3	(protection (scheme or propert\$3 or parameter or type)) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S316	21	(backup path) with (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S317	0	(backup path) with (protection near\$3 parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S318	131	(backup path) with (protection )	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S319	1	(genera\$5 or configur\$5) with (backup path) with (parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S320	164	(genera\$5 or configur\$5) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S321	24	(genera\$5 or configur\$5) with (backup path) and (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S322	0	(genera\$5 or configur\$5) with (backup path) with (base or according)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S323	0	(genera\$5 or configur\$5) with (backup path) with ("base" or "according")	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55

S324	181	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S325	17	(genera\$5 or configur\$5 or setup) with (backup path) not S320	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S326	2654	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S327	4210	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S328	366	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S329	6	(709/220).ccls. and S328	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
S330	1	"20050226215"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/01/15 11:55
S331	2	"20060045028"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/01/15 11:55
S332	490	(pseudowire or pseudo-wire or pseudo wire)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S333	15	(pseudowire or pseudo-wire or pseudo wire) and (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S334	102856	standby path with priority	US-PGPUB; USPAT	OR	ON	2010/07/13 18:19
S335	0	standby path with priority	US-PGPUB; USPAT	ADJ	ON	2010/07/13 18:19
S336	3	standby path with priority	US-PGPUB; USPAT	ADJ	ON	2010/07/13 18:19
S337	7	switch\$5 with priority same (standby path)	US-PGPUB; USPAT	ADJ	ON	2010/07/13 18:26
S338	1	preempt with existing traffic	US-PGPUB; USPAT	ADJ	ON	2010/07/13 18:29
S339	199	preempt with traffic	US-PGPUB; USPAT	ADJ	ON	2010/07/13 18:31
S340	67	preempt with traffic with priority	US-PGPUB; USPAT	ADJ	ON	2010/07/13 18:31
S341	64	LDp same acknow\$11	US-PGPUB; USPAT	ADJ	ON	2010/07/14 14:22
S342	56	LDp same acknow\$11 and @ad< "20050216"	US-PGPUB; USPAT	ADJ	ON	2010/07/14 14:22
S343	12	LDp same acknow\$11 and @ad< "20050216" and (label distribution protocol)	US-PGPUB; USPAT	ADJ	ON	2010/07/14 14:23
S344	0	Ping near2 pan and pseduo \$5	US-PGPUB; USPAT	ADJ	ON	2010/07/14 15:36
S345	0	(Ping near2 pan).inv. and pseduo\$5	US-PGPUB; USPAT	ADJ	ON	2010/07/14 15:37

S346	7	(Ping near2 pan).inv. and pseudo\$5	US-PGPUB; USPAT	ADJ	ON	2010/07/14 15:37
S347	183	((pseudo-wire) or (pseudowire)) and LDP	US-PGPUB; USPAT	ADJ	ON	2010/07/14 15:50
S348	57	((pseudo-wire) or (pseudowire)) and LDP and @ad<"20050216"	US-PGPUB; USPAT	ADJ	ON	2010/07/14 15:50
S349	2	S348 and backup path	US-PGPUB; USPAT	ADJ	ON	2010/07/14 15:56
S350	38	LDP with protection	US-PGPUB; USPAT	ADJ	ON	2010/07/14 16:21
S351	29	LDP with protection and @ad<"20050216"	US-PGPUB; USPAT	ADJ	ON	2010/07/14 16:24
S352	3	"7,385,920"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/07/15 14:26
S353	1	"20060109786".pn.	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/07/15 15:13
S354	8	protection with domain type	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/07/16 16:06
S355	4	protection and domain type and LDP	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/07/16 16:07
S356	60	single-hop and multi-hop and backup	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/07/16 16:15
S357	3	single-hop and multi-hop and backup path	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/07/16 16:15
S358	6	multi-hop pseudowire	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/07/16 16:23
S359	0	(domain type) with pseudowire	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/07/16 16:28
S360	0	(hop type) with pseudowire	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/07/16 16:28
S361	28	(\$7hop) with pseudowire	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/07/16 16:28
S363	0	pseudowire and tele	US-PGPUB; USPAT	OR	ON	2010/07/16 16:52
S368	0	(pseudowire or pseudo-wire) and iniliz\$5	US-PGPUB; USPAT	OR	ON	2010/07/16 16:52
S387	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:52

S390	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:52
S393	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S389) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:52
S394	441	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/07/16 16:52
S395	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S394) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:52
S397	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S396) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:52
S400	3980648	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2010/07/16 16:52
S401	1566705	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:52
S402	4233012	(alter\$7 or backup or standby)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:52
S403	29916257	@ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:52
S404	7292265	(pick\$5 or select\$5 or choos \$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:52
S409	214	(S402 near\$ S400) with config\$7 with (primary near\$ S402)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:52
S416	4327773	(send\$7 or transmit\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:52
S417	0	(source node) with S416 with (config\$7 near\$ parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:52

S421	3928	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S422	0	(ack or acknowledgement) same (config\$7 parameter \$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S423	28099	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S424	56105	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S427	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S428	10997	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S433	4339	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S434	2738	(protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S438	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S439	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S443	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/07/16 16:53
S444	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/07/16 16:53
S446	0	receiv\$7 acknowledgement (parameter accept\$5 destination node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/07/16 16:53

S452	4	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/07/16 16:53
S454	0	(domain type) with (single- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S455	0	(domain type) with (single near5 hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S459	0	field with indica\$7 with ((single-hop) same (multi- hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S462	312	(single-hop) same (multi- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S466	0	S464 same (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S468	0	parameter with indicat\$5 same ((single hop) or (single-hop)) same ((multi- hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S469	0	(field or parameter) with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S470	0	(field or parameter) with (show\$3 or indicat\$5) same ((single hop) or (single- hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S471	0	(field or parameter) with (domain type) same ((multi- hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S472	72	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S475	3193	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53

S480	424	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S481	338	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S486	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S488	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S491	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 netowrk)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S493	0	(paramete or field or bit) with indicat\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53
S497	77123	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/07/16 16:53
S510	0	(backup path) with (protection near3 parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/07/16 16:53
S515	0	(genera\$5 or configur\$5) with (backup path) with (base or according)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/07/16 16:53
S516	0	(genera\$5 or configur\$5) with (backup path) with ("base" or "according")	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/07/16 16:53
S517	204	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/07/16 16:53
S519	2921	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/07/16 16:53
S520	4629	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2010/07/16 16:53
S521	441	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/07/16 16:53
S522	8	(709/220).ccls. and S521	US-PGPUB; USPAT	OR	ON	2010/07/16 16:53
S525	589	(pseudowire or pseudo-wire or pseudo wire)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/07/16 16:53
S532	0	(pseudowire or pseudo-wire) and initiliz\$5	US-PGPUB; USPAT	OR	ON	2010/07/16 16:53



S551	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:53
S554	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:53
S557	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S553) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:53
S558	441	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/07/16 16:53
S559	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S558) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:53
S561	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S560) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/07/16 16:53
S564	3980648	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2010/07/16 16:53
S565	1566705	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/07/16 16:53

**EAST Search History (Interference)**

&lt; This search history is empty &gt;

**7/ 16/ 2010 4:59:59 PM****C:\ Documents and Settings\ sliu3\ My Documents\ EAST\ Workspaces\ 11354569.wsp**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No.: 11/354,569	Confirmation No.: 6912
Applicant: Ping Pan	Group Art Unit: 2472
Filing Date: February 14, 2006	Examiner: Liu, Siming
Docket No.: 002.P045	
Customer No.: 65638	
For: PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE	<b>AMENDMENT AND RESPONSE TO OFFICE ACTION MAILED JANUARY 22, 2010  SUBMITTED THROUGH EFS-WEB</b>

**AMENDMENT**

Dear Sir:

In response to the Office Action mailed January 22, 2010, Applicant respectfully requests that the following amendment be made part of the official record in the above captioned case.

Amendments to the Claims begin on page 2 of this paper.

Amendments to the Drawings are described on page 6 of this paper and are also included in both attached replacement sheets and annotated sheets.

Remarks begin at page 7 of this paper.

**AMENDMENTS TO THE CLAIMS:**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims:**

1. (Currently Amended) A method of providing protection to network traffic, comprising:
  - sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;
  - receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; ~~and~~
  - accepting the Pseudowire protection configuration parameter by the destination node;
  - using the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter[[]] ; and
  - determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.
2. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.
3. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.
4. (Original) A method as recited in claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.
5. (Currently Amended) A method as recited in claim 1, wherein the ~~Pseudowire protection property configuration parameter~~ further includes at least one of a domain type, a protection type or a protection scheme.

6. (Canceled)

7. (Canceled)

8. (Canceled)

9. (Canceled)

10. (Original) A method as recited in claim 1, wherein the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP).

11. (Currently Amended) A system for providing protection to network traffic, comprising:  
a processor configured to:

send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;

receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and

~~in the even that~~ accept the Pseudowire protection configuration parameter has been accepted by the destination node[[,]] ;

use the standby Pseudowire; ~~wherein the standby Pseudowire that~~ is configured based at least in part on the Pseudowire protection configuration parameter[[.]] ; and

determine whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

12. (Original) A system as recited in Claim 11, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.

13. (Currently Amended) A system as recited in Claim 11, wherein the ~~Pseudowire~~ protection

property configuration parameter further includes at least one of a domain type, a protection type or a protection scheme.

14. (Canceled)

15. (Canceled)

16. (Canceled).

17. (Currently Amended) A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for:

    sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;

    receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; ~~and in the even that accept the Pseudowire protection configuration parameter has been accepted by the destination node[[,]] ;~~

    using the standby Pseudowire; ~~wherein the standby Pseudowire that is configured based at least in part on the Pseudowire protection configuration parameter[[.]] ; and~~

determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.

18. (Currently Amended) A computer program product as recited in claim 17, wherein the Pseudowire protection property configuration parameter further includes at least one of a domain type, a protection type or a protection scheme.

19. (Canceled)

20. (Canceled)

21. (Canceled).

22. (New) A method as recited in claim 5, wherein the domain type indicates whether the Pseudowire is configured in a single-hop environment where the Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the Pseudowire includes a plurality of nodes coupled to several carrier networks.

23. (New) A method as recited in claim 5, wherein the protection scheme indicates at least one of the following:

a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;

a 1:1 protection scheme, wherein one Pseudowire is used to protect another Pseudowire;

a 1:N protection scheme, wherein one Pseudowire is used to protect N other Pseudowires;

or

an M:N protection scheme, wherein M Pseudowires are used to protect N other Pseudowires.

24. (New) A system as recited in claim 13, wherein the domain type indicates whether the Pseudowire is configured in a single-hop environment where the Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the Pseudowire includes a plurality of nodes coupled to several carrier networks.

25. (New) A system as recited in claim 13, wherein the protection scheme indicates at least one of the following:

a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;

a 1:1 protection scheme, wherein one Pseudowire is used to protect another Pseudowire;

a 1:N protection scheme, wherein one Pseudowire is used to protect N other Pseudowires;

or

an M:N protection scheme, wherein M Pseudowires are used to protect N other Pseudowires.

26. (New) A computer product as recited in claim 18, wherein the domain type indicates

whether the Pseudowire is configured in a single-hop environment where the Pseudowire includes a plurality of nodes coupled to a same carrier network, or a multi-hop environment where the Pseudowire includes a plurality of nodes coupled to several carrier networks.

27. (New) A computer product as recited in claim 18, wherein the protection scheme indicates at least one of the following:

- a 1+1 protection scheme, wherein the same traffic is sent over two Pseudowires;
- a 1:1 protection scheme, wherein one Pseudowire is used to protect another Pseudowire;
- a 1:N protection scheme, wherein one Pseudowire is used to protect N other Pseudowires;

or

an M:N protection scheme, wherein M Pseudowires are used to protect N other Pseudowires.

**AMENDMENTS TO THE DRAWINGS:**

The attached sheets of drawings include changes to FIG. 4. Changes to FIG. 4 are presented to make FIG. 4 consistent with its description in the Specification. No new matter has been introduced.

Attachments:            Replacement Sheets for FIG. 4.  
                                 Annotated Sheet Showing changes to FIG. 4.



**REMARKS**

The above-referenced patent application has been reviewed in light of the Office Action mailed **January 22, 2010** (the “Action”). In the Action, claims 10-21 were objected to due to informalities. Claims 1-4, 7, 11-12, 15, 17 and 20 were rejected under 35 U.S.C. § 103(a) as being unpatentable over US 2003/0117950 to Huang (“Huang”), in view of US 2006/0047851 to Voit *et al.* (“Voit”), US 2004/0133692 to Blanchet *et al.* (“Blanchet”) and US 2006/0018252 to Sridhar *et al.* (“Sridhar”). Claims 5, 13 and 18 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Huang in view of Voit, Blanchet and Sridhar and further in view of US 2006/0046658 to Cruz *et al.* (“Cruz”). Claims 6, 14 and 19 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Huang in view of Voit, Blanchet and Sridhar and further in view of US 6,574,477 to Rathude (“Rathude”). Finally, claims 8-9, 16 and 21 were rejected under 35 U.S.C. §103(a) as being unpatentable over Huang, Voit, Blanchet and Sridhar and further in view of US 7,200,104 to Saleh *et al.* (“Saleh”).

**Current Status of Claims:**

With this amendment, claims 1-5, 10-13, 17, 18 and 22-27 are pending. Applicant offers to amend claims 1, 5, 11, 13, 17 and 18, as presented above. Applicant has canceled claims 6-9, 14-16 and 19-21. Applicant has also added new claims 22-27, as presented above. No new matter has been introduced.

**Objection to claims 10-21:**

Claims 10-21 (either canceled or amended) are included with this Response. Thus, Applicant requests that the objection to pending claims 10-13, 17 and 18 be withdrawn.

**Rejection of claims 1-4, 11-12 and 17 under 35 U.S.C. § 103(a):**

A portion of claim 1, as currently amended, recites:

“sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire, the protection property including a priority for the standby Pseudowire;...

*determining whether to preempt existing traffic on the standby Pseudowire, wherein the determination is based, at least in part, on the priority for the standby Pseudowire.”*

Emphasis added.

The portion of amended claim 1, as presented above, includes elements of canceled claim 9. As admitted in the Action for the rejection of claim 9, Huang in view of Voit, Blanchet and Sridhar does not disclose, “determining whether to preempt existing traffic on the standby Pseudowire, the determination being based at least in part on a priority associated with the standby Pseudowire”. (See Action, page 12). The Action relies on Saleh to remedy the admitted deficiencies in Huang, Voit, Blanchet and Sridhar. However, Saleh does not disclose at least the above-emphasized portions of claim 1.

The Action points to Saleh’s description of a QoS as being “equivalent to priority.” But Saleh merely discloses that QoS criteria can be used for **selecting what nodes can be a part of a virtual path or VP**. (See Col. 3, lines 3-8). Therefore, Saleh’s description of the use of QoS criteria for node selection does not disclose, “*determining whether to preempt existing traffic on the standby Pseudowire, ... based, at least in part, on a priority associated with the standby Pseudowire*” (emphasis added).

Saleh also discloses a priority or Class of Service (CoS) that determines a VP’s relative priority for performance and restoration in the event of a failure. (See Col. 3, lines 37-41). Saleh further describes provisioning two distinct physical paths for a given VP, one for a primary path and a second for a secondary path. (See Col. 4, lines 19-24). However, Saleh merely describes **assigning a CoS to the VP in general** and does not describe separately assigning a CoS to the primary and secondary paths. Furthermore, the secondary path is dedicated to the given VP for restoration purposes and is **only used in case of a failure in the primary path**. (See Col. 4, lines 27-29). Since a CoS type of priority is assigned to only the VP in general and/or the secondary path is **used only in case of a failure** of a primary path, Saleh does not disclose, *determining whether to preempt existing traffic on the standby Pseudowire, ... based, at least in part, on a priority associated with the standby Pseudowire*” (emphasis added).

For at least the above-mentioned portions of claim 1, Huang, Voit, Blanchet and Sridhar in view of Saleh do not support a *prima facie* 35 U.S.C. §103(a) rejection of claim 1. Applicant requests that the rejection of claim 1 be withdrawn.

Independent claims 11 and 17 include similar elements to those mentioned above for claim 1. Additionally, claims 2-4 and 12 depend from one of claims 1, 11 or 17. Thus, Applicant requests that the 35 U.S.C. §103(a) rejections of 2-4, 11, 12 and also be withdrawn.

**Rejection of claims 5, 13 and 18 under 35 U.S.C. § 103(a):**

Claims 5, 13 and 18 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Huang, in view of Voit, Blanchet and Sridhar and further in view of Cruz. Claims 5, 13 and 18 depend from claims 1, 11 and 17, respectively. As a result, for the same reasons mentioned above for claim 1, Huang, Voit, Blanchet and Sridhar do not support a *prima facie* 35 U.S.C. §103(a) rejection of claims 5, 13 and 18. Also, Cruz does not cure the above-stated deficiencies of Huang, Voit, Blanchet and Sridhar. Thus, Applicant requests that the 35 U.S.C. §103(a) rejections of claims 5, 13 and 18 be withdrawn.

**Rejection of claims 6, 8, 9, 14, 16, 19 and 21 under 35 U.S.C. § 103(a):**

Claims 6, 8, 9, 14, 16, 19 and 21 have been canceled, so this rejection is moot.

**Conclusion:**

Applicant respectfully submits that claims 1-5, 10-13, 17, 18 and 22-27 are in condition for allowance and such action is earnestly solicited. *The Examiner is respectfully requested to contact the undersigned by telephone at (503) 551-9442 if it is believed that such contact would further the examination of the present application.*

Respectfully submitted,

Date: 4/21/2010

by: /Ted A. Crawford/Reg. No. 50,610/  
Ted A. Crawford  
Reg. No. 50,610



Annotated Sheet  
Application No. 11/354,569  
Attorney Docket No. 002.P045  
Sheet 1 of 1

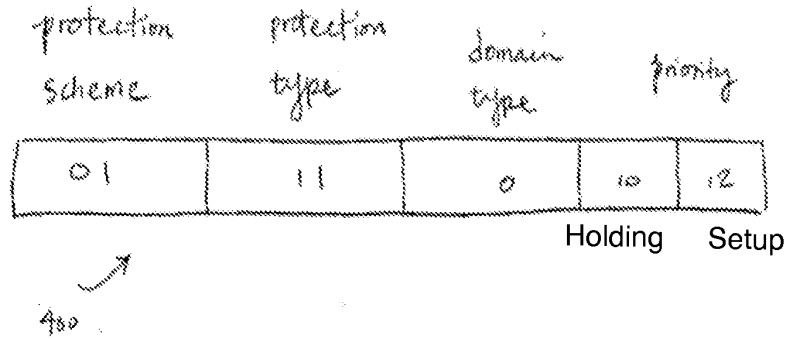
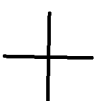
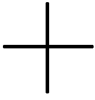


FIG. 4





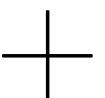
Replacement Sheet  
Application No. 11/354,569  
Attorney Docket No. 002.P045  
Sheet 1 of 1

Protection Scheme	Protection Type	Domain Type	Priority	
01	11	0	10	12

Holding    Setup

400 

FIG. 4



## Electronic Acknowledgement Receipt

<b>EFS ID:</b>	7461963
<b>Application Number:</b>	11354569
<b>International Application Number:</b>	
<b>Confirmation Number:</b>	6912
<b>Title of Invention:</b>	Pseudowire protection using a standby pseudowire
<b>First Named Inventor/Applicant Name:</b>	Ping Pan
<b>Customer Number:</b>	65638
<b>Filer:</b>	Ted A. Crawford/Lindsey Hunt
<b>Filer Authorized By:</b>	Ted A. Crawford
<b>Attorney Docket Number:</b>	002.P045
<b>Receipt Date:</b>	21-APR-2010
<b>Filing Date:</b>	14-FEB-2006
<b>Time Stamp:</b>	18:27:30
<b>Application Type:</b>	Utility under 35 USC 111(a)

### Payment information:

Submitted with Payment	no
------------------------	----

### File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Applicant Arguments/Remarks Made in an Amendment	Response_to_OA_11_354569_002_P045.pdf	145825 <small>15b483d4c5ec0934a667a123efcd98417e84f8ad</small>	no	11

### Warnings:

### Information:

2	Drawings-only black and white line drawings	Annotated_Sheet_FIG_4_002_P045.pdf	52251 3d4d779268006c9e1be3c3dd34fc35ebabcb52b	no	1
<b>Warnings:</b>					
<b>Information:</b>					
3	Drawings-only black and white line drawings	Replacement_Sheet_Fig_4_002_P045.pdf	35671 84450a8ac336126a2ded739b8e83ab8cd79ea8e3	no	1
<b>Warnings:</b>					
<b>Information:</b>					
<b>Total Files Size (in bytes):</b>			233747		
<p><b>This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.</b></p> <p><b><u>New Applications Under 35 U.S.C. 111</u></b>  If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.</p> <p><b><u>National Stage of an International Application under 35 U.S.C. 371</u></b>  If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.</p> <p><b><u>New International Application Filed with the USPTO as a Receiving Office</u></b>  If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.</p>					



Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

<b>PATENT APPLICATION FEE DETERMINATION RECORD</b> Substitute for Form PTO-875					Application or Docket Number <b>11/354,569</b>		Filing Date <b>02/14/2006</b>		<input type="checkbox"/> To be Mailed			
<b>APPLICATION AS FILED – PART I</b>												
(Column 1)			(Column 2)		SMALL ENTITY <input type="checkbox"/>		OR			OTHER THAN SMALL ENTITY		
FOR		NUMBER FILED	NUMBER EXTRA		RATE (\$)	FEE (\$)	OR		RATE (\$)	FEE (\$)		
<input type="checkbox"/> BASIC FEE <small>(37 CFR 1.16(a), (b), or (c))</small>		N/A	N/A		N/A				N/A			
<input type="checkbox"/> SEARCH FEE <small>(37 CFR 1.16(k), (l), or (m))</small>		N/A	N/A		N/A		N/A					
<input type="checkbox"/> EXAMINATION FEE <small>(37 CFR 1.16(o), (p), or (q))</small>		N/A	N/A		N/A		N/A					
TOTAL CLAIMS <small>(37 CFR 1.16(i))</small>		minus 20 =	*		X \$ =		OR		X \$ =			
INDEPENDENT CLAIMS <small>(37 CFR 1.16(h))</small>		minus 3 =	*		X \$ =				X \$ =			
<input type="checkbox"/> APPLICATION SIZE FEE <small>(37 CFR 1.16(s))</small>		If the specification and drawings exceed 100 sheets of paper, the application size fee due is \$250 (\$125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).										
<input type="checkbox"/> MULTIPLE DEPENDENT CLAIM PRESENT <small>(37 CFR 1.16(j))</small>												
* If the difference in column 1 is less than zero, enter "0" in column 2.												
<b>APPLICATION AS AMENDED – PART II</b>												
(Column 1)			(Column 2)		(Column 3)		SMALL ENTITY		OR		OTHER THAN SMALL ENTITY	
AMENDMENT	<b>04/21/2010</b>		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE (\$)	ADDITIONAL FEE (\$)		RATE (\$)	ADDITIONAL FEE (\$)	
	Total <small>(37 CFR 1.16(j))</small>		* 17	Minus	** 21	= 0	X \$ =		OR		X \$52=	0
	Independent <small>(37 CFR 1.16(h))</small>		* 3	Minus	***3	= 0	X \$ =		OR		X \$220=	0
	<input type="checkbox"/> Application Size Fee <small>(37 CFR 1.16(s))</small>											
	<input type="checkbox"/> FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM <small>(37 CFR 1.16(j))</small>											
TOTAL ADD'L FEE						TOTAL ADD'L FEE						
						<b>0</b>						
AMENDMENT			CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE (\$)	ADDITIONAL FEE (\$)		RATE (\$)	ADDITIONAL FEE (\$)	
	Total <small>(37 CFR 1.16(j))</small>		*	Minus	**	=	X \$ =		OR		X \$ =	
	Independent <small>(37 CFR 1.16(h))</small>		*	Minus	***	=	X \$ =		OR		X \$ =	
	<input type="checkbox"/> Application Size Fee <small>(37 CFR 1.16(s))</small>											
	<input type="checkbox"/> FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM <small>(37 CFR 1.16(j))</small>											
TOTAL ADD'L FEE						TOTAL ADD'L FEE						
* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.										OR		
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".												
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".												
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.										Legal Instrument Examiner: /SANDRA F. GARNETT/		

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NUMBER	PATENT NUMBER	GROUP ART UNIT	FILE WRAPPER LOCATION
11/354,569		2472	



**Correspondence Address/Fee Address Change**

The following fields have been set to Customer Number 65638 on 03/18/2010

- Correspondence Address
- Power of Attorney Address

The address of record for Customer Number 65638 is:

65638  
OMIKRON IP LAW GROUP  
16325 Boones Ferry Rd.  
SUITE 204  
LAKE OSWEGO, OR 97035



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
11/354,569	02/14/2006	Ping Pan	HAMMP008

65638  
OMIKRON IP LAW GROUP  
16325 Boones Ferry Rd.  
SUITE 204  
LAKE OSWEGO, OR 97035

**CONFIRMATION NO. 6912**  
**POA ACCEPTANCE LETTER**



Date Mailed: 03/17/2010

**NOTICE OF ACCEPTANCE OF POWER OF ATTORNEY**

This is in response to the Power of Attorney filed 03/05/2010.

The Power of Attorney in this application is accepted. Correspondence in this application will be mailed to the above address as provided by 37 CFR 1.33.

/tha/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

MAR 05 2010

PTO/SB/08 (11-08)

Approved for use through 11/30/2011, OMB 0051-0038  
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**POWER OF ATTORNEY TO PROSECUTE APPLICATIONS BEFORE THE USPTO**

I hereby revoke all previous powers of attorney given in the application identified in the attached statement under 37 CFR 3.73(b).

I hereby appoint:

Practitioner associated with the Customer Number: 65638

OR  
 Practitioner(s) named below (if more than ten patent practitioners are to be named, then a customer number must be used):

Name	Registration Number	Name	Registration Number

as attorney(s) or agent(s) to represent the undersigned before the United States Patent and Trademark Office (USPTO) in connection with any and all patent applications assigned only to the undersigned according to the USPTO assignment records or assignment documents attached to this form in accordance with 37 CFR 3.73(b).

Please change the correspondence address for the application identified in the attached statement under 37 CFR 3.73(b) to:

The address associated with Customer Number: 65638

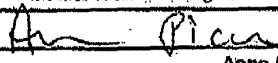
OR

Firm or Individual Name	
Address	
City	State Zip
Country	
Telephone	Email

Assignee Name and Address:  
Brixham Solutions Ltd.  
OMC Chambers, Wickhams Cay 1, Road Town,  
Tortola, British Virgin Islands

A copy of this form, together with a statement under 37 CFR 3.73(b) (Form PTO/SB/08 or equivalent) is required to be filed in each application in which this form is used. The statement under 37 CFR 3.73(b) may be completed by one of the practitioners appointed in this form if the appointed practitioner is authorized to act on behalf of the assignee, and must identify the application in which this Power of Attorney is to be filed.

SIGNATURE of Assignee of Record  
The individual whose signature and title is supplied below is authorized to act on behalf of the assignee

Signature		Date
Name	Anna C. Picchi	Telephone
Title	Authorized Person	

This collection of information is required by 37 CFR 1.31, 1.32 and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

*If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.*

MAR 05 2010

PTO/SB/06 (07-09)  
Approved for use through 07/31/2012. OMB 0831-0031  
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**STATEMENT UNDER 37 CFR 3.73(b)**

Applicant/Patent Owner: Brixham Solutions Ltd.

Application No./Patent No.: 11/354,569

Filed/Issue Date: 2/14/2006

Titled: PSEUDOWIRE PROTECTION

Brixham Solutions Ltd., a corporation  
(Name of Assignee) (Type of Assignee, e.g., corporation, partnership, university, government agency, etc.)

states that it is:

- 1.  the assignee of the entire right, title, and interest in;
- 2.  an assignee of less than the entire right, title, and interest in (The extent (by percentage) of its ownership interest is \_\_\_\_\_ %); or
- 3.  the assignee of an undivided interest in the entirety of (a complete assignment from one of the joint inventors was made) the patent application/patent identified above, by virtue of either:

A.  An assignment from the inventor(s) of the patent application/patent identified above. The assignment was recorded in the United States Patent and Trademark Office at Reel \_\_\_\_\_, Frame \_\_\_\_\_, or for which a copy therefore is attached.

OR

B.  A chain of title from the inventor(s), of the patent application/patent identified above, to the current assignee as follows:

1. From: Pan, Ping To: Hammerhead Systems

The document was recorded in the United States Patent and Trademark Office at Reel 017613, Frame 0722, or for which a copy thereof is attached.

2. From: Hammerhead Systems, Inc. To: Brixham Solutions Ltd.

The document was recorded in the United States Patent and Trademark Office at Reel 023810, Frame 0916, or for which a copy thereof is attached.

3. From: \_\_\_\_\_ To: \_\_\_\_\_

The document was recorded in the United States Patent and Trademark Office at Reel \_\_\_\_\_, Frame \_\_\_\_\_, or for which a copy thereof is attached.

Additional documents in the chain of title are listed on a supplemental sheet(s).

As required by 37 CFR 3.73(b)(1)(i), the documentary evidence of the chain of title from the original owner to the assignee was, or concurrently is being, submitted for recordation pursuant to 37 CFR 3.11.

[NOTE: A separate copy (i.e., a true copy of the original assignment document(s)) must be submitted to Assignment Division in accordance with 37 CFR Part 3, to record the assignment in the records of the USPTO. See MPEP 302.08]

The undersigned (whose title is supplied below) is authorized to act on behalf of the assignee.

Ted Crawford  
Signature

3/5/10  
Date

Ted A. Crawford

Authorized Attorney

Printed or Typed Name

Title

This collection of information is required by 37 CFR 3.73(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.



16325 Boones Ferry Road Suite 204  
Lake Oswego, Oregon 97035  
P: 503.719.9473  
F: 503.305.6760

RECEIVED  
CENTRAL FAX CENTER  
MAR 05 2010

OmikronIPLawGroup

Facsimile Cover Sheet

To:	USPTO
Fax Number:	(571) 273-8300
Phone Number:	
From:	Lindsey Hunt
Fax Number:	(503) 305-6760
Phone Number:	(503) 719-9473
Date:	March 5, 2010
Pages including this cover page:	3

Please find attached the following:

Power of Attorney to Prosecute Application Before the USPTO appointing Customer Number 65638

Statement Under 37 CFR 3.73(b) for Application No. 11/354,569

CERTIFICATE OF FACSIMILE	
I hereby certify that this correspondence (along with any paper referred to as being attached or enclosed) is being facsimile transmitted to the United States Patent and Trademark Office (Fax No. (571) 273-8300 on <u>March 5, 2010</u> .)	
Date	<u>March 5, 2010</u>
Signature	<u><i>Lindsey Hunt</i></u>
Typed or printed name of person signing Certificate	
<u>Lindsey Hunt</u>	

This information is intended to be for the use of the individual or entity named on this transmittal sheet. If you are not the intended recipient, be aware that any disclosure, copying, distribution or use of the contents of this faxed information is prohibited. If you have received this facsimile in error, please notify the sender by telephone immediately so that arrangements can be made for the retrieval of the original document at no cost to you.

MAR 05 2010

PTO/SB/80 (11-08)

Approved for use through 11/30/2011. OMB 0851-0033  
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**POWER OF ATTORNEY TO PROSECUTE APPLICATIONS BEFORE THE USPTO**

I hereby revoke all previous powers of attorney given in the application identified in the attached statement under 37 CFR 3.73(b).

I hereby appoint:

Practitioners associated with the Customer Number: 65638

OR  
Practitioner(s) named below (if more than ten patent practitioners are to be named, then a customer number must be used):

Name	Registration Number	Name	Registration Number

as attorney(s) or agent(s) to represent the undersigned before the United States Patent and Trademark Office (USPTO) in connection with any and all patent applications assigned only to the undersigned according to the USPTO assignment records or assignment documents attached to this form in accordance with 37 CFR 3.73(b).

Please change the correspondence address for the application identified in the attached statement under 37 CFR 3.73(b) to:

The address associated with Customer Number: 65638

Firm or Individual Name

Address

City State Zip

County

Telephone Email

Assignee Name and Address:

Brixham Solutions Ltd.  
OMC Chambers, Wickhams Cay 1, Road Town,  
Tortola, British Virgin Islands

A copy of this form, together with a statement under 37 CFR 3.73(b) (Form PTO/SB/88 or equivalent) is required to be filed in each application in which this form is used. The statement under 37 CFR 3.73(b) may be completed by one of the practitioners appointed in this form if the appointed practitioner is authorized to act on behalf of the assignee, and must identify the application in which this Power of Attorney is to be filed.

SIGNATURE of Assignee of Record  
The individual whose signature and title is supplied below is authorized to act on behalf of the assignee

Signature		Date
Name	Anna C. Pichi	Telephone
Title	Authorized Person	

This collection of information is required by 37 CFR 1.31, 1.32 and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/SB/96 (07-09)

Approved for use through 07/31/2012. OMB 0651-0031  
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**STATEMENT UNDER 37 CFR 3.73(b)**

Applicant/Patent Owner: Brixham Solutions Ltd.

Application No./Patent No.: 11/354,569 Filed/Issue Date: 2/14/2006

Titled: PSEUDOWIRE PROTECTION

Brixham Solutions Ltd., a corporation  
(Name of Assignee) (Type of Assignee, e.g., corporation, partnership, university, government agency, etc.)

states that it is:

- 1.  the assignee of the entire right, title, and interest in;
- 2.  an assignee of less than the entire right, title, and interest in (The extent (by percentage) of its ownership interest is \_\_\_\_\_ %); or
- 3.  the assignee of an undivided interest in the entirety of (a complete assignment from one of the joint inventors was made)

the patent application/patent identified above, by virtue of either:

A.  An assignment from the inventor(s) of the patent application/patent identified above. The assignment was recorded in the United States Patent and Trademark Office at Reel \_\_\_\_\_, Frame \_\_\_\_\_, or for which a copy therefore is attached.

OR

B.  A chain of title from the inventor(s), of the patent application/patent identified above, to the current assignee as follows:

1. From: Pan, Ping To: Hammerhead Systems

The document was recorded in the United States Patent and Trademark Office at Reel 017613, Frame 0722, or for which a copy thereof is attached.

2. From: Hammerhead Systems, Inc. To: Brixham Solutions Ltd.

The document was recorded in the United States Patent and Trademark Office at Reel 023810, Frame 0916, or for which a copy thereof is attached.

3. From: \_\_\_\_\_ To: \_\_\_\_\_

The document was recorded in the United States Patent and Trademark Office at Reel \_\_\_\_\_, Frame \_\_\_\_\_, or for which a copy thereof is attached.

Additional documents in the chain of title are listed on a supplemental sheet(s).

As required by 37 CFR 3.73(b)(1)(i), the documentary evidence of the chain of title from the original owner to the assignee was, or concurrently is being, submitted for recordation pursuant to 37 CFR 3.11.

[NOTE: A separate copy (i.e., a true copy of the original assignment document(s)) must be submitted to Assignment Division in accordance with 37 CFR Part 3, to record the assignment in the records of the USPTO. See MPEP 302.08]

The undersigned (whose title is supplied below) is authorized to act on behalf of the assignee.

Ted Crawford  
Signature

3/5/10  
Date

Ted A. Crawford

Authorized Attorney

Printed or Typed Name

Title

This collection of information is required by 37 CFR 3.73(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 36 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.





16325 Boones Ferry Road Suite 204  
 Lake Oswego, Oregon 97035  
 P: 503.719.9473  
 F: 503.305.6760

**OmikronIPLawGroup**

RECEIVED  
 CENTRAL FAX CENTER  
 MAR 05 2010

### Facsimile Cover Sheet

To:	USPTO
Fax Number:	(571) 273-8300
Phone Number:	
From:	Lindsey Hunt
Fax Number:	(503) 305-6760
Phone Number:	(503) 719-9473
Date:	March 5, 2010
Pages including this cover page:	3

Please find attached the following:

**Power of Attorney to Prosecute Application Before the USPTO appointing  
 Customer Number 65638**

**Statement Under 37 CFR 3.73(b) for Application No. 11/354,569**

CERTIFICATE OF FACSIMILE	
I hereby certify that this correspondence (along with any paper referred to as being attached or enclosed) is being facsimile transmitted to the United States Patent and Trademark Office (Fax No. (571) 273-8300 on <u>March 5, 2010</u> .	
<u>March 5, 2010</u> Date	<u>Lindsey Hunt</u> Signature
	<u>Lindsey Hunt</u> Typed or printed name of person signing Certificate

This information is intended to be for the use of the individual or entity named on this transmittal sheet. If you are not the intended recipient, be aware that any disclosure, copying, distribution or use of the contents of this faxed information is prohibited. If you have received this facsimile in error, please notify the sender by telephone immediately so that arrangements can be made for the retrieval of the original document at no cost to you.

PAGE 1/3 \* RCVD AT 3/5/2010 7:15:50 PM [Eastern Standard Time] \* SVR:USPTO-EFXXF-6/34 \* DNIS:2738300 \* CSID:5033056760 \* DURATION (mm-ss):01-56

JUNIPER Exhibit 1003

App. 3, pg. 167

'652 File History 167



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	HAMMP008	6912
21912	7590	01/22/2010	EXAMINER	
VAN PELT, YI & JAMES LLP 10050 N. FOOTHILL BLVD #200 CUPERTINO, CA 95014			LIU, SIMING	
			ART UNIT	PAPER NUMBER
			2472	
			MAIL DATE	DELIVERY MODE
			01/22/2010	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b> 11/354,569	<b>Applicant(s)</b> PAN, PING	
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 2472	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1)  Responsive to communication(s) filed on 12/22/2009.
- 2a)  This action is **FINAL**.
- 2b)  This action is non-final.
- 3)  Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4)  Claim(s) 1-21 is/are pending in the application.
  - 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5)  Claim(s) \_\_\_\_\_ is/are allowed.
- 6)  Claim(s) 1-21 is/are rejected.
- 7)  Claim(s) \_\_\_\_\_ is/are objected to.
- 8)  Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9)  The specification is objected to by the Examiner.
- 10)  The drawing(s) filed on \_\_\_\_\_ is/are: a)  accepted or b)  objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11)  The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12)  Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    - a)  All    b)  Some \*    c)  None of:
    - 1.  Certified copies of the priority documents have been received.
    - 2.  Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
    - 3.  Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1)  Notice of References Cited (PTO-892)
- 2)  Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3)  Information Disclosure Statement(s) (PTO/SB/08)  
 Paper No(s)/Mail Date \_\_\_\_\_
- 4)  Interview Summary (PTO-413)  
 Paper No(s)/Mail Date. \_\_\_\_\_
- 5)  Notice of Informal Patent Application
- 6)  Other: \_\_\_\_\_

## DETAILED ACTION

### *Response to Arguments*

1. Applicant's arguments filed on 12/22/2009 have been fully considered but they are not persuasive. First, applicant's argument is not directed the most recent office action. Applicant still response to the first office action sent on 11/20/2008. However, a more recent office action was sent out on 06/22/2009.

2. Regarding to applicant's argument that it's not obvious to combine the references to obtain the claimed invention. Applicant' specially mentioned the invention solves long felt but unresolved needs and failure of others to provide adequate Pseudowire protection. However, reference Voit teaches Pseudowire and Pseudowire protection. Therefore, the concept of provide adepquate Pseudowire protection is not new giving the reference Voit, US 2006/0047851 A1.

### *Claim Objections*

1. Claims 10-21 are objected to because of the following informalities: Claims 11-21 are not received with the RCE application. Only a portion of claim 10 is received. However, base on the submitted claims 1-9. Applicant seems want to keep the original

claims and forgot to copy and paste the previous presented claims. The claims 10-21 are examined based on the assumption that those claims are still same as the previous presented claim set filled on 02/24/2009. Appropriate correction is required.

***Claim Rejections - 35 USC § 103***

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

2. Claims 1-4, 7, 11-12, 15, 17, 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang US 2003/0117950 A1, in view of Voit US 2006/0047851 A1, further in view of Blanchet US 2004/0133692 A1, further in view of Sridhar US 2006/0018252 A1.

3. Regarding claim 1, Huang teaches a method of providing protection to network traffic (Huang, page 2, [0015], lines 1-6), comprising:  
sending (Huang, page 2, [0016], right column, line 1: "receiving a request to set up". There must be sending, thus receiving can happen) a ... protection configuration parameter (Huang, page 2, [0016], right column, lines 2-4: "the request specifying a required protection bandwidth for the label switched path segment", "required protection bandwidth" can be considered as a protection configuration parameter) for configuring a standby ... between a source node and a destination node (Huang, page 2, [0016], right

column, lines 4-5: "and determining a backup route to the tail end node"), ...; receiving a ... configuration acknowledgement indicating whether the ... protection configuration parameter has been accepted by the destination node; and accepting the ... protection configuration parameter by the destination node, using the standby ... (Huang, page 2, [0016], right column, lines 6-14: In response of the request of setting up a label switched path segment over a direct connection between two nodes, a backup route is also being determined); using the standby path that is configured based at least in part on the ... protection configuration parameter (Huang, page 2, [0016], right column, lines 8-12: "The method also includes signaling to reserve the required protection bandwidth along the backup route, receiving confirmation of reservation of the required protection bandwidth and generating a backup connection map", required protection bandwidth as a configuration parameter is a major factor in the backup path forming).

Huang doesn't expressly teach Pseudowire and Pseudowire protection.

Voit teaches Pseudowire (Voit, page 2, [0011], lines 2-7) and Pseudowire protection (Voit, page 4, [0046], lines 1-3: "a network topology is provided with redundant pseudowire connections ...").

At the time of the invention was made, it would have been obvious to a person of ordinary skill in the art to implement Pseudowire as a type of network service in the system disclosed by Huang in order to increase communication security since Voit teaches that PWs can provide point-to-point connectivity and are similar to virtual private link. Voit also teaches providing data traffic protection for primary Pseudowire

path (Voit, page 4, [0046], lines 1-3). Both Huang and Voit are in the same field of endeavor (network transfer) and are directed to the same problem sought to be solved (data traffic protection).

Huang in view of Voit doesn't expressly teach that receiving a configuration acknowledgement indicating whether the configuration parameter has been accepted by the destination node.

Blanchet teaches that receiving a configuration acknowledgement indicating whether the configuration parameter has been accepted by the destination node (Blanchet, page 4, [0035], lines 2-4).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the system to send an ACK indicating the acceptance of the configuration parameters in the system disclosed by Huang in view of Voit in order to makes the system more reliable. Both Huang in view of Voit and Blanchet are in the same field of endeavor (Network transfer).

Huang in view of Voit and Blanchet doesn't expressly teach that the configuration parameter indicating a protection property associated with the standby link.

Sridhar teaches that the configuration parameter indicating a protection property associated with the standby link (Sridar, [0031], lines 3-10: it is noted the backup path setup is based on one of the three protection schemes, the parameter indicate the type of protection scheme is considered as the configuration parameters).

At the time of the invention was made, it would have been obvious to a person of ordinary skill in the art to configure a backup path based on the configuration

parameters which indicate a property associated with the standby link in the system disclosed by Huang in view of Voit and Blanchet in order increase the efficiency of the system. Since not all primary paths have the same risk, it enables the system to weight the risk and assign proper resources to each primary path for protection. Both Huang in view of Voit, Blanchet and Sridhar are in the same field of endeavor (Network transfer) and are directed to the same problem sought to be solved (data traffic protection).

1. Regarding claims 2, 12, Huang in view of Voit, Blanchet and Sridhar further teaches the standby (Huang, page 2, [0016], right column, lines 5: "backup" is equivalent to standby in the context) Pseudowire (VOIT, [0002], lines 1-2) is configured to provide protection to at least one primary (Huang, page 1, [0008], lines 2-4) Pseudowire.

2. Regarding claim 3, Huang in view of Voit, Blanchet and Sridhar further teaches the standby Pseudowire is configured to provide protection to at least one primary Pseudowire (Huang, page 1, [0008], lines 2-4), and in the event that the primary Pseudowire (VOIT, [0002], lines 1-2) fails to transfer network traffic (Huang, page 2, [0010], line 5: "when a fault is discovered in a single link between two nodes along a path"), switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire (Huang, page 2, [0010], lines 9-10: "switches the traffic that was using the connection to the alternate path", the limitation "Pseudowire" has been discussed).



3. Regarding claim 4, Huang in view of Voit, Blanchet and Sridhar further teaches the standby Pseudowire is dynamically selected from a plurality of connections (Huang, page 4, [0040], lines 12-14: "The backup route may, for instance, be selected from a table of routes that have been pre-computed to connect the head end node 102A to the tail end node 102B"; page 4, [0040], right column, lines 1-2: "a backup route can be determined instantaneously by the head end node 102A given information about the current state of the network 100").

4. Regarding claims 7, 15, 20, Huang in view of Voit, Blanchet and Sridhar further teaches the Pseudowire protection configuration parameter includes a protection scheme (Sridhar, [0031], lines 3-10: it is noted the backup path setup is based on one of the three protection schemes, the parameter indicate the type of protection scheme is considered as the configuration parameters).

5. Regarding claim 10, Huang in view of Voit, Blanchet and Sridhar further teaches the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP) (Huang, page 1, [0005], right column, last 3 lines).

6. Regarding claim 11, Huang in view of Voit, Blanchet and Sridhar teaches a system for providing protection to network traffic, comprising:  
a processor (Blanchet, Fig. 2, element 50: It's inherent, since all computers have at

least one processor) configured to:

send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire; and

receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, use the standby Pseudowire; wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter; and a memory coupled to the processor, configured to provide the processor with instructions (Blanchet, Fig. 2, element 50: it's inherent since all computers have a memory coupled to a processor, and provide instructions with processor). (All of the remaining limitations have been discussed in claim 1)

7. Regarding claim 17, Huang in view of Voit, Blanchet and Sridhar teaches a computer program product (Huang, page 2, [0013]: "the 'gold' level of service protection" is a computer program product) for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions (Blanchet, Fig. 2, element 50: it's inherent, since all computers have memory and can give

computer instructions) for:

sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire;

receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire; wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter (All of the remaining limitations have been discussed in claim 1).

8. Claims 5, 13, 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over under 35 U.S.C. 103(a) as being unpatentable over Huang, in view of Voit, Blanchet and Sridhar, further in view of Cruz, US 2006/0046658 A1.

9. Regarding claims 5, 13, 18, Huang in view of VOIT, Blanchet and Sridhar as applied in claim above teaches a method as recited in Claim 1, wherein the Pseudowire (VOIT, [0002], lines 1-2) protection configuration parameter (Huang, page 2, [0016], right column, lines 2-4: “the request specifying a required protection bandwidth for the label switched path segment”, “required protection bandwidth” is equivalent to a protection configuration parameter) includes ...

Huang in view of Voit, Blanchet and Sridhar doesn't expressly teach that a domain type.

Cruz teaches a domain type (Cruz, page 1, para 0017, line 2: According to the specification of the application, domain type is about whether the network is either multi-hop or single hop).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include domain type. The reason is that by including domain type in the configuration parameter, it would be more accurate to select a desired standby path, given that you have more information about the network. The method of change the configuration parameter by including the domain type of Huang in view of VOIT, Blanchet and Sridhar was within the ordinary ability of one of ordinary skill in the art based on the teachings of Cruz.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, VOIT, Blanchet, Sridhar and Cruz to obtain the invention as specified in claims 5, 13, 18.

10. Claims 6, 14, 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang in view of VOIT, Blanchet and Sridhar, further in view of Rathunde, US 6,574,477 B1.

11. Regarding claims 6, 14, 19, Huang in view of Voit, Blanchet and Sridhar teaches a method as recited in Claim 1, wherein the Pseudowire protection configuration parameter (previous discussed) includes ...

Huang in view of Voit, Blanchet and Sridhar doesn't expressly teach that a protection type.

Rathunde teaches a protection type (Rathunde, col 9, line 3: "type of standby mode", according to the specification of the application, protection type just means what type of standby mode).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include a protection type. The reason is that by including protection type in the configuration parameter, it would be more accurate to select a desire standby path, given that you have more information about the network. The method of change the configuration parameter by including the protection type of Huang in view of VOIT and Blanchet was within the ordinary ability of one of ordinary skill in the art based on the teachings of Rathunde.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, VOIT, Blanchet and Rathunde to obtain the invention as specified in claims 6, 14, 19.

12. Claims 8-9, 16, 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang, Voit, Blanchet and Sridhar, further in view of Saleh, US 7,200,104 B2.

13. Regarding claims 8, 16, 21, Huang in view of Voit, Blanchet and Sridhar teaches a method as recited in Claim 1, wherein the Pseudowire protection configuration parameter (previous discussed) includes a ...

Huang in view of Voit, Blanchet and Sridhar doesn't expressly teach that a priority.

Saleh teaches a priority (Saleh, col 3, line 38: "restoration priority level").

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include priority. The reason is that by including domain type in the configuration parameter, it would be more accurate to select a desired standby path, given that you have more information about the network. The method of change the configuration parameter by including the domain type of Huang in view of Voit and Blanchet was within the ordinary ability of one of ordinary skill in the art based on the teachings of Saleh.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, Voit, Blanchet and Saleh to obtain the invention as specified in claims 8, 16, 21.

14. Regarding claim 9, Huang in view of Voit, Blanchet and Sridhar teaches a method as recited in Claim 1, further including determining whether to preempt existing traffic on the standby Pseudowire, the determination being based at least in part on a priority associated with the standby Pseudowire (Saleh, col 3, lines 3-8: QoS is equivalent to priority).

**Conclusion**

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIMING LIU whose telephone number is (571)270-3859. The examiner can normally be reached on Monday-Friday 8:30am-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, William Trost can be reached on 571-272-7872. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/William Trost/  
Supervisory Patent Examiner, Art  
Unit 2472

/S. L./  
Examiner, Art Unit 2472

<b>Notice of References Cited</b>	Application/Control No. 11/354,569	Applicant(s)/Patent Under Reexamination PAN, PING	
	Examiner SIMING LIU	Art Unit 2472	Page 1 of 1

**U.S. PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A US-2006/0018252	01-2006	Sridhar et al.	370/216
*	B US-2006/0047851	03-2006	Voit et al.	709/239
	C US-			
	D US-			
	E US-			
	F US-			
	G US-			
	H US-			
	I US-			
	J US-			
	K US-			
	L US-			
	M US-			

**FOREIGN PATENT DOCUMENTS**


*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N				
	O				
	P				
	Q				
	R				
	S				
	T				

**NON-PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)				
	U				
	V				
	W				
	X				

\*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)  
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.




<b>Index of Claims</b>  	<b>Application/Control No.</b> 11354569	<b>Applicant(s)/Patent Under Reexamination</b> PAN, PING
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 2472

✓	<b>Rejected</b>	-	<b>Cancelled</b>	N	<b>Non-Elected</b>	A	<b>Appeal</b>
=	<b>Allowed</b>	÷	<b>Restricted</b>	I	<b>Interference</b>	O	<b>Objected</b>

Claims renumbered in the same order as presented by applicant
  CPA
  T.D.
  R.1.47

CLAIM		DATE								
Final	Original	10/30/2008	06/17/2009	01/14/2010						
	1	✓	✓	✓						
	2	✓	✓	✓						
	3	✓	✓	✓						
	4	✓	✓	✓						
	5	✓	✓	✓						
	6	✓	✓	✓						
	7	✓	✓	✓						
	8	✓	✓	✓						
	9	✓	✓	✓						
	10	✓	✓	✓						
	11	✓	✓	✓						
	12	✓	✓	✓						
	13	✓	✓	✓						
	14	✓	✓	✓						
	15	✓	✓	✓						
	16	✓	✓	✓						
	17	✓	✓	✓						
	18	✓	✓	✓						
	19	✓	✓	✓						
	20	✓	✓	✓						
	21	✓	✓	✓						

<b>Search Notes</b>  	<b>Application/Control No.</b>  11354569	<b>Applicant(s)/Patent Under Reexamination</b>  PAN, PING
	<b>Examiner</b>  SIMING LIU	<b>Art Unit</b>  2472

SEARCHED			
Class	Subclass	Date	Examiner
370	216, 225, 228	10/30/2008	/SL/
709	220	10/30/2008	/SL/
above	update search	6/17/2009	/SL/
update search	ABOVE	1/14/2010	/SL/

SEARCH NOTES		
Search Notes	Date	Examiner
East Class search	11/10/2008 update 6/17/2009	/SL/
Palm inventor name search	10/30/2008 update 6/17/2009	/SL/
Consulted 101 issues with Peng, John	11/10/2008	/SL/
update search: ABOVE	1/14/2010	/SL/

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner

--	--

## EAST Search History

## EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	199	pseudowire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L3	366	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L4	11	L3 with protection	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L5	4	L3 with protection and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L6	1	"20040223498".pn.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L7	0	(pseudowire or pseudo-wire) and initiliz\$5	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L8	209	(pseudowire or pseudo-wire) and initi\$5	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L9	71	(pseudowire or pseudo-wire) and initi\$5 and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L10	2654	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L11	14	(370/216,225,228).ccls. and L3	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L12	6	(709/220).ccls. and L3	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L13	35	((PING) near2 (PAN)).INV.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L14	2	((PING) near2 (PAN)).INV. and pseudowire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L15	2	((PING) near2 (PAN)).INV. and (pseudowire).clm.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L16	100	L3 and (primary)	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L17	30	L3 and (primary) and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L18	106	L3 and (config\$7) and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L19	4	TDM pseudowire	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L20	189	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55

L21	62	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L22	14	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (standby or backup)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L23	12	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (primary or main) and (secondly or backup or standby)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L24	68	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L25	22	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L26	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L27	32	(pseudowire or (pseudo wire) or pseudo-wire) and ((config\$7) same (destination near5 node))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L28	4	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with acknowledgement)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L29	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L30	12	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L31	441	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge))	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L32	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (L28) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55

L33	366	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L34	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (L33) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L35	462	pseudowire or pseudo-wire or (pseudo wire)	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L36	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (L35) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L37	246	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L38	41	L35 and initialization	US-PGPUB; USPAT	ADJ	ON	2010/01/15 11:55
L39	3810399	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2010/01/15 11:55
L40	1494211	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L41	4045518	(alter\$7 or backup or standby)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L42	29886744	@ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L43	7016429	(pick\$5 or select\$5 or choos\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L44	420	(L39 near7 L40) with (L43 near7 L41 near7 L39) and L42	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L45	246	(L39 near7 L40) with (L43 near7 L41 near7 L39) and L42 and (priority or bandwidth)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L46	28	(L39 near7 L40) with (L43 near7 L41 near7 L39) same (priority or bandwidth or parameter) and L42	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

L47	2735	L41 with config\$7 with (primary near7 L41)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L48	188	(L41 near5 L39) with config \$7 with (primary near7 L41)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L49	207	(L41 near5 L39) with config \$7 with (primary near7 L39)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L50	119	(L41 near5 L39) with config \$7 with (primary near7 L39) and L42	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L51	7	09/859166	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L52	39	(restoration scheme) and ("1:N")	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L53	5	(restoration scheme) and (priority) and (standby mode)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L54	22	(restoration scheme) and (priority) and (config\$7 near5 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L55	4138472	(send\$7 or transmit\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L56	0	(source node) with L55 with (config\$7 near3 parameter \$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L57	7	(source) with L55 with (config\$7 near3 parameter \$1) with (destin\$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L58	675	(source) with L55 with (parameter\$1) with (destin \$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L59	62	(source node) with L55 with (parameter\$1) with (destin \$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

L60	3563	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L61	0	(ack or acknowledgement) same (config\$7 parameter \$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L62	25761	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L63	55233	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L64	31	(ack or acknowledgement) and (restoration scheme) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L65	153	L55 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L66	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L67	10895	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L68	797	handshaking and @ad<"20050214" and (L39 with L40)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L69	2	"6553034".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L70	114	(virtual path) and ((protection or restoration) near5 scheme) and priority	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L71	106	(virtual path) and ((protection or restoration) near5 scheme) and priority and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L72	4061	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

L73	2702	(protection or restoration) near5 parameter and @ad< "20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L74	7	((protection or restoration) near5 parameter) with (L55) and (destin\$7 near3 node) and @ad< "20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L75	27	((protection or restoration) near5 parameter) and (handshaking) and @ad< "20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L76	76	((protection or restoration) near5 parameter) and (destination node) and @ad< "20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L77	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad< "20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L78	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad< "20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L79	16	((protection or restoration or config\$7) near5 parameter) with (destination node) and @ad< "20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L80	4	((protection or restoration or config\$7) near2 parameter) with (destination node) and @ad< "20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L81	693	handshaking and @ad< "20050214" and (config\$7 parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L82	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L83	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L84	325	receiv\$7 acknowledgement destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L85	0	receiv\$7 acknowledgement (parameter accept\$5 destination node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55



L86	5	receiv\$7 acknowledgement accept\$3 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L87	8	receiv\$7 acknowledgement parameter accept\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L88	2	"20030117950".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L89	999	(domain type) with (parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L90	4	(parameter) near5 includ\$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L91	4	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L92	71	(parameter\$1) with (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2010/01/15 11:55
L95	71	(domain type) with parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L96	258	(single-hop) same (multi- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L97	11	(single-hop) same (multi- hop) same (parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L98	0	field with indica\$7 with ((single-hop) same (multi- hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L99	197	field with indica\$7 with (topology)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L100	10	field with indica\$7 with (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

L101	258	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L102	18	(field or parameter) same ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L103	415	((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L104	35	(parameter or field) same L103	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L106	150	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L110	0	(field or parameter) with (domain type) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L111	71	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L112	15	(field or parameter) with (indicat\$5 or show\$5) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L113	1	(protection type) and (standby path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L114	2982	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L115	292	(hot and cold) same standby and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L116	52	(hot and cold) and (parameter with standby) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L117	21	(field with indicat\$5 with (standby mode)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

L118	725	config\$9 with (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L119	421	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L120	336	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L121	203	config\$9 near3 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L122	7	config\$9 near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L123	9	type with (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L124	4	type near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L125	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L126	150	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L127	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L128	142	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L129	14	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) same (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L130	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 netowrk)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55

L131	193	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 network)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L132	0	(paramete or field or bit) with indicat\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L133	13	(paramete or field or bit) with indicat\$7 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2010/01/15 11:55
L134	1	(protection near3 properties) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2010/01/15 11:55
L135	8	(protection near3 (parameter or propert\$5)) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2010/01/15 11:55
L136	70110	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/01/15 11:55
L137	16	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2010/01/15 11:55
L138	37	(pseudowire or pseudo-wire) and (standby)	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L139	5	(pseudowire or pseudo-wire) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L140	12	(protection scheme) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L141	1	(protection scheme) with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L142	1	(protection (type or property)) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L143	11	((protection (type or property)) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L144	20	(( (type or property) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L145	1	(protection scheme) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L146	2	(protection scheme) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55

L147	3	(protection (scheme or property or parameter or type)) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L148	21	(backup path) with (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L149	0	(backup path) with (protection near parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L150	131	(backup path) with (protection )	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L151	1	(general or configuration) with (backup path) with (parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L152	164	(general or configuration) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L153	24	(general or configuration) with (backup path) and (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L154	0	(general or configuration) with (backup path) with (base or according)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L155	0	(general or configuration) with (backup path) with ("base" or "according")	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L156	181	(general or configuration or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L157	17	(general or configuration or setup) with (backup path) not L152	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
L158	2654	(370/216,225,228).cls.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L159	4210	(709/220).cls.	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L160	366	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L161	6	(709/220).cls. and L160	US-PGPUB; USPAT	OR	ON	2010/01/15 11:55
L162	1	"20050226215"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/01/15 11:55
L163	2	"20060045028"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2010/01/15 11:55
L164	490	(pseudowire or pseudo-wire or pseudo wire)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55

L165	15	(pseudowire or pseudo-wire or pseudo wire) and (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2010/01/15 11:55
S5	115	pseudowire	US-PGPUB; USPAT	OR	ON	2008/10/08 12:53
S6	0	pseudowire and tele	US-PGPUB; USPAT	OR	ON	2008/10/08 12:53
S7	217	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2008/10/08 13:07
S8	9	S7 with protection	US-PGPUB; USPAT	OR	ON	2008/10/08 13:08
S9	4	S7 with protection and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 13:09
S10	1	"20040223498".pn.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:07
S11	0	(pseudowire or pseudo-wire) and iniliz\$5	US-PGPUB; USPAT	OR	ON	2008/10/08 14:14
S12	133	(pseudowire or pseudo-wire) and initi\$5	US-PGPUB; USPAT	OR	ON	2008/10/08 14:15
S13	51	(pseudowire or pseudo-wire) and initi\$5 and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:15
S14	2193	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:17
S15	6	(370/216,225,228).ccls. and S7	US-PGPUB; USPAT	OR	ON	2008/10/08 14:23
S16	3	(709/220).ccls. and S7	US-PGPUB; USPAT	OR	ON	2008/10/08 14:26
S17	31	((PING) near2 (PAN)).INV.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:32
S18	2	((PING) near2 (PAN)).INV. and pseudowire	US-PGPUB; USPAT	OR	ON	2008/10/08 14:33
S19	2	((PING) near2 (PAN)).INV. and (pseudowire).clm.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:33
S20	66	S7 and (primary)	US-PGPUB; USPAT	OR	ON	2008/10/08 14:38
S21	23	S7 and (primary) and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:39
S22	75	S7 and (config\$7) and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:44
S23	2	TDM pseudowire	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:09
S24	106	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol))	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:16
S26	43	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad< "20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:17

S27	11	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (standby or backup)	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:19
S28	9	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (primary or main) and (secondly or backup or standby)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 09:00
S29	43	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:34
S30	14	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:35
S31	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:38
S32	18	(pseudowire or (pseudo wire) or pseudo-wire) and ((config\$7) same (destination near5 node))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:38
S33	1	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with acknowledgement)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:41
S34	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:44
S35	8	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:44
S36	370	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:47
S37	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S33) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:48
S38	218	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2008/10/09 10:49

S39	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S38) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:49
S40	275	pseudowire or pseudo-wire or (pseudo wire)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:49
S41	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S40) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:50
S42	233	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:50
S43	27	S40 and initialization	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:54
S44	3434011	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2008/10/29 09:26
S46	1334019	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:27
S47	3640734	(alter\$7 or backup or standby)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:28
S48	29788561	@ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:28
S49	6386553	(pick\$5 or select\$5 or choos\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:30
S50	409	(S44 near7 S46) with (S49 near7 S47 near7 S44) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:31
S51	238	(S44 near7 S46) with (S49 near7 S47 near7 S44) and S48 and (priority or bandwidth)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:38
S52	25	(S44 near7 S46) with (S49 near7 S47 near7 S44) same (priority or bandwidth or parameter) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:43
S53	2289	S47 with config\$7 with (primary near7 S47)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:49



S54	159	(S47 near5 S44) with config \$7 with (primary near7 S47)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:54
S55	175	(S47 near5 S44) with config \$7 with (primary near7 S44)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:54
S56	111	(S47 near5 S44) with config \$7 with (primary near7 S44) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:56
S57	7	09/859166	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 10:26
S58	33	(restoration scheme) and ("1:N")	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 10:50
S59	4	(restoration scheme) and (priority) and (standby mode)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 12:42
S60	18	(restoration scheme) and (priority) and (config\$7 near5 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:27
S61	3706723	(send\$7 or transmit\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:30
S62	0	(source node) with S61 with (config\$7 near3 parameter \$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:31
S63	5	(source) with S61 with (config\$7 near3 parameter \$1) with (destin\$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:32
S64	569	(source) with S61 with (parameter\$1) with (destin \$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:33
S65	54	(source node) with S61 with (parameter\$1) with (destin \$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:33
S66	2959	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:53

S67	0	(ack or acknowledgement) same (config\$7 parameter \$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:53
S68	20807	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:54
S69	52906	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:54
S70	29	(ack or acknowledgement) and (restoration scheme) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:55
S71	137	S61 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:19
S72	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:29
S73	10549	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:29
S74	759	handshaking and @ad<"20050214" and (S44 with S46)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:30
S75	2	"6553034".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:32
S76	108	(virtual path) and ((protection or restoration) near5 scheme) and priority	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:34
S77	103	(virtual path) and ((protection or restoration) near5 scheme) and priority and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:34
S78	3479	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:51
S79	2628	(protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:52

S80	6	((protection or restoration) near5 parameter) with (S61) and (destin\$7 near3 node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:53
S81	26	((protection or restoration) near5 parameter) and (handshaking) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:55
S82	73	((protection or restoration) near5 parameter) and (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:04
S83	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:04
S84	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S85	15	((protection or restoration or config\$7) near5 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S86	4	((protection or restoration or config\$7) near2 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S87	651	handshaking and @ad<"20050214" and (config\$7 parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:10
S88	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:27
S89	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:27
S90	276	receiv\$7 acknowledgement destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:28
S91	0	receiv\$7 acknowledgement (parameter accept\$5 destination node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:40
S92	5	receiv\$7 acknowledgement accept\$3 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:45

S93	7	receiv\$7 acknowledgement parameter accept\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:48
S94	2	"20030117950".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:20
S95	771	(domain type) with (parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:25
S96	3	(parameter) near5 includ\$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:26
S97	3	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:27
S98	64	(parameter\$1) with (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:27
S99	0	(domain type) with (single- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:28
S100	0	(domain type) with (single near5 hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:29
S101	64	(domain type) with parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:29
S102	179	(single-hop) same (multi- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:32
S103	9	(single-hop) same (multi- hop) same (parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:33
S104	0	field with indica\$7 with ((single-hop) same (multi- hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:34
S105	147	field with indica\$7 with (topology)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:34

S106	6	field with indica\$7 with (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:36
S107	179	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:37
S108	10	((field or parameter) same ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:37
S109	283	((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:40
S111	21	(parameter or field) same S109	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:41
S112	0	S109 same (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:44
S113	134	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:46
S114	0	parameter with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:48
S115	0	(field or parameter) with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:49
S116	0	(field or parameter) with (show\$3 or indicat\$5) same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:49
S117	0	(field or parameter) with (domain type) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S118	68	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50

S119	14	(field or parameter) with (indicat\$5 or show\$5) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S120	1	(protection type) and (standby path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:44
S121	2636	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:46
S122	283	(hot and cold) same standby and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:47
S123	51	(hot and cold) and (parameter with standby) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:48
S124	20	(field with indicat\$5 with (standby mode)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:49
S126	696	config\$9 with (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:50
S127	406	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:51
S128	324	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:51
S129	194	config\$9 near3 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:52
S130	7	config\$9 near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:54
S131	9	type with (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:01
S132	4	type near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:01

S133	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:38
S135	134	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:39
S136	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:43
S137	126	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:43
S138	11	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) same (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:44
S139	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 netowrk)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:53
S140	136	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 network)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:53
S141	0	(paramete or field or bit) with indicat\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:55
S142	2	(paramete or field or bit) with indicat\$7 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:55
S143	1	(protection near3 properties) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2009/05/02 14:27
S144	7	(protection near3 (parameter or proptert\$5)) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2009/05/02 14:28
S145	61963	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2009/05/02 15:12
S146	15	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2009/05/02 15:13
S147	25	(pseudowire or pseudo-wire) and (standby)	US-PGPUB; USPAT	OR	ON	2009/05/02 15:20

S148	3	(pseudowire or pseudo-wire) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 15:20
S149	12	(protection scheme) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:04
S150	1	(protection scheme) with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:06
S151	1	(protection (type or property)) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:22
S152	10	((protection (type or property)) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:23
S153	19	(( (type or property) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:26
S154	1	(protection scheme) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:27
S155	2	(protection scheme) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:27
S156	3	(protection (scheme or propert\$3 or parameter or type)) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:29
S157	20	(backup path) with (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:22
S158	0	(backup path) with (protection near\$3 parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:35
S159	121	(backup path) with (protection )	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:35
S160	1	(genera\$5 or configur\$5) with (backup path) with (parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:36
S161	144	(genera\$5 or configur\$5) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:37
S162	22	(genera\$5 or configur\$5) with (backup path) and (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:38
S163	0	(genera\$5 or configur\$5) with (backup path) with (base or according)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:42
S164	0	(genera\$5 or configur\$5) with (backup path) with ("base" or "according")	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:42



S165	159	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:43
S166	15	(genera\$5 or configur\$5 or setup) with (backup path) not S161	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:43
S168	2421	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S169	3849	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S170	291	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S171	5	(709/220).ccls. and S170	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S172	1	"20050226215"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2009/06/17 16:27
S173	2	"20060045028"	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2009/06/17 16:30
S174	397	(pseudowire or pseudo-wire or pseudo wire)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 16:41
S175	13	(pseudowire or pseudo-wire or pseudo wire) and (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 16:41

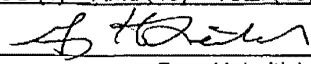
**EAST Search History (Interference)**

&lt; This search history is empty &gt;

**1/15/2010 12:10:14 PM****C:\Documents and Settings\slu3\My Documents\EAST Workspaces\11354569.wsp**

RECEIVED  
CENTRAL FAX CENTER  
DEC 22 2009  
PATENT  
ATTORNEY DOCKET  
NO. HAMMP0008

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants:	Ping Pan	)	CERTIFICATE OF FACSIMILE TRANSMISSION
		)	
Serial No.:	11/354,569	)	The undersigned hereby certifies that this
		)	document is being facsimile transmitted to the fax
Filed:	February 14, 2006	)	number and date given below.
		)	
Title:	PSEUDO-WIRE PROTECTION	)	Date Transmitted: <u>December 21, 2009</u>
	USING A STANDBY	)	Facsimile Number: <u>571-273-8300</u>
	PSEUDOWIRE	)	No. of Pages: <u>Cov (1) + RCE (2) + Pet.Ext.Time</u>
		)	<u>(1) + CC (1) + Amd (10) = Total (15)</u>
Group Art Unit:	2416	)	
Examiner:	LIU, Siming	)	By: 
		)	Greg H. Leitich

REQUEST FOR CONTINUED EXAMINATION

Mail Stop RCE  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

Applicant hereby requests continued examination, in accordance with 37 C.F.R. §1.114, for the above-identified application.

TIME REQUEST IS BEING MADE

1. Prosecution on this application being closed (under either appeal, final action, notice of allowance, or other prosecution closing action), this request is being submitted prior to the earliest of:

- i.  Abandonment of the application.
- ii.  Payment of issue fee, or
- Issue fee has been paid but a petition under 37 C.F.R. §1.313 has been granted.
- iii.  Filing of a notice of appeal to the U.S. Court of Appeals for the Federal Circuit under 35 U.S.C. §1.313, or commencement of a civil action under 35 U.S.C. 145 or 146, unless the appeal or civil action is terminated.
- iv.  A decision on appeal to the Board of Patent Appeals & Interferences (this RCE is to be treated as a request to withdraw the appeal and to reopen prosecution of the application) - A notice is being separately sent to the Board of Patent Appeals & Interferences that this RCE is being filed.

Serial No.: 11/486,432

Attorney Docket No.: HAMMP0013

SUBMISSIONS AND ENCLOSURES

- 2. Enclosed herewith is/are:
  - i.  **A Petition for Extension of Time for three (3) month(s).**
  - ii.  **The enclosed Preliminary Amendment.**
  - iii.  Please enter the previously unentered Amendment faxed 18 June 2008. A copy  is or  is not provided herein.
  - iv.  An Information Disclosure Statement (37 C.F.R. §1.98) with PTO-1449 and \_\_\_\_\_ copies of references.
  - v.  New arguments or new evidence in support of patentability.
  - vi.  Other: \_\_\_\_\_

FEE FOR REQUEST REQUIRED BY 37 C.F.R. §1.17(e)

- 3.  Filing fee has been calculated as shown below (small entity) after entering the previous amendment and/or currently submitted amendment as may be applicable:

	Current Claims Pending Minus Highest Number Previously Paid For	No. Extra	Rate	Fees
Total Claims	21-21	= 0	x \$52 =	\$ 0.00
Indep. Claims	3- 3	= 0	x \$110 =	\$ 0.00
<input type="checkbox"/> Multiple Dependent Claims Present			+ \$390 =	\$ 0.00
Basic filing fee				\$ 810.00
Total				\$ 810.00

- 4.  A credit card payment form including the amount of the fee is enclosed.

PLEASE MAIL CORRESPONDENCE TO:

Greg H. Leitich  
P.O. Box 3255  
Austin TX 78703

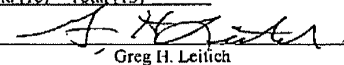
Respectfully submitted,

  
 \_\_\_\_\_  
 Greg H. Leitich, Reg. No. 39,745  
 Attorney(s) for Applicant(s)  
 Direct Dial: 512-469-0063

December 21, 2009

RECEIVED  
CENTRAL FAX CENTER  
DEC 22 2009

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants:	Ping Pan	)	CERTIFICATE OF FACSIMILE TRANSMISSION
Serial No.:	11/354,569	)	The undersigned hereby certifies that this document is being facsimile transmitted to the fax number and date given below.
Filed:	February 14, 2006	)	Date Transmitted: <u>December 21, 2009</u>
Title:	PSEUDO-WIRE PROTECTION USING A STANDBY PSEUDOWIRE	)	Facsimile Number: <u>571-273-8300</u>
Docket No.	HAMMP008	)	No. of Pages: <u>Cov. (1) + RCF. (2) + Pet. Ext. Time (1) + CC (1)</u>
Group Art Unit:	2416	)	+ Amd (10) = Total (15)
Examiner:	LIU, Siming	)	By: <u></u> Greg H. Leitich

**Petition For Extension of Time**

Box Amendment FEE  
Commissioner of Patents and Trademarks  
Washington, DC 20231

Dear Sir:

This is a request under the provisions of 37 CFR 1.136(a) to extend the period for filing a response to the office action in the above identified application.

The requested extension and appropriate fee are as follows:

<input type="checkbox"/>	One month (37 CFR 1.17(a)(1))	\$
<input type="checkbox"/>	Two months (37 CFR 1.17(a)(2))	\$
<input checked="" type="checkbox"/>	Three months (37 CFR 1.17(a)(3))	\$ 1110.00
<input type="checkbox"/>	Four months (37 CFR 1.17(a)(4))	\$
<input type="checkbox"/>	Five months (37 CFR 1.17(a)(5))	\$

Applicant is a small entity under 37 CFR 1.9 and 1.27, therefore the fee amount shown above is reduced by one-half, and the resulting fee is: \$ \_.

A small entity statement under 37 CFR 1.27:

is enclosed.

has already been filed in this application.

A check in the amount of the fee is enclosed

A credit card payment form including the amount of the fee is enclosed

The Commissioner has already been authorized to charge fees in this application to a Deposit Account.

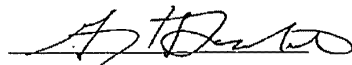
I am the:

assignee of record of the entire interest.

applicant.


attorney or agent of record.

attorney or agent under 37 CFR 1.34(a). Reg. No: 39,745

Date December 21, 2009 Name: Greg H. Leitich Signature: 

RECEIVED  
CENTRAL FAX CENTER  
DEC 22 2009

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants:	Ping Pan	)	CERTIFICATE OF FACSIMILE TRANSMISSION
Serial No.:	11/354,569	)	The undersigned hereby certifies that this document is
Filed:	February 14, 2006	)	being facsimile transmitted to the fax number and date
Title:	PSEUDO-WIRE PROTECTION USING A STANDBY PSEUDOWIRE	)	given below.
		)	Date Transmitted: <u>December 21, 2009</u>
		)	Facsimile Number: <u>571-273-8300</u>
		)	No. of Pages: <u>Cov (1) + RCE (2) + Pet. Ext. Time (1)</u>
		)	<u>+ CC (1) + Amd (10) = Total (15)</u>
Group Art Unit:	2416	)	By: 
Examiner:	LIU, Siming	)	Greg H. Leitich

AMENDMENT

Dear Sir:

In response to the final Official Action mailed June 22, 2009, applicants respectfully request that the following amendment be made part of the official record in the above captioned case. The applicant also submits a petition for a 3 month extension of time and submits a request for continued examination (RCE).

Amendments to the claims begin on page 2 of this paper.

Remarks begin at page 6 of this paper.

Application Serial No. 11/354,569

Patent

**AMENDMENTS TO THE CLAIMS:**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims:**

1. (Currently Amended) A method of providing protection to network traffic, comprising:
  - sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire;
  - receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and
  - ~~in the event that~~ accepting the Pseudowire protection configuration parameter ~~has been accepted~~ by the destination node[[,]] ;
  - using the standby Pseudowire; ~~wherein the standby Pseudowire~~ that is configured based at least in part on the Pseudowire protection configuration parameter.
2. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.
3. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.
4. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.
5. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration

Application Serial No. 11/354,569

Patent

parameter includes a domain type.

6. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a protection type.
7. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a protection scheme.
8. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a priority.
9. (Original) A method as recited in Claim 1, further including determining whether to preempt existing traffic on the standby Pseudowire, the determination being based at least in part on a priority associated with the standby Pseudowire.
10. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration

PAGE 8/8 \* RCVD AT 12/21/2009 12:14:00 PM [Eastern Standard Time] \* SVR:USPTO-EFXXRF-6/15 \* DNIS:2738300 \* CSID:512 469 0023 \* DURATION (mm-ss):02-46

RECEIVED  
CENTRAL FAX CENTER  
DEC 22 2009

GREG H. LEITICH

P.O. BOX 3255  
AUSTIN, TEXAS 78764

TELEPHONE: (512) 469-0063  
FACSIMILE: (512) 469-0023  
E-MAIL: leitich@sbcglobal.net

FAX TRANSMITTAL COVER SHEET


TO: Examiner LIU  
USPTO 571-273-8300  
Group Art Unit 2416

DATE: December 21, 2009

FROM: Greg H. Leitich

MATTER: 11/354,569

TOTAL NUMBER OF PAGES (INCLUDING THIS ONE): 15

Applicants:	Ping Pan	)	CERTIFICATE OF FACSIMILE TRANSMISSION
Serial No.:	11/354,569	)	The undersigned hereby certifies that this document is
Filed:	February 14, 2006	)	being facsimile transmitted to the fax number and date
Title:	PSEUDO-WIRE PROTECTION USING A STANDBY PSEUDOWIRE	)	given below.
Group Art Unit:	2416	)	Date Transmitted: <u>December 21, 2009</u>
Examiner:	LIU, Siming	)	Facsimile Number: <u>571-273-8300</u>
		)	No. of Pages: <u>Cov (1) + RCE (2) +</u>
		)	<u>Pet, Exl. Time (1) + CC (1) + Amd (10) = Total (15)</u>
		))	By:  Greg H. Leitich

This facsimile and the information it contains is intended to be a confidential communication only to the person or entity to whom it is addressed. If you have received this facsimile in error, please notify us immediately by calling us collect at the above-listed number(s).



Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

<b>PATENT APPLICATION FEE DETERMINATION RECORD</b> Substitute for Form PTO-875					Application or Docket Number <b>11/354,569</b>		Filing Date <b>02/14/2006</b>		<input type="checkbox"/> To be Mailed			
<b>APPLICATION AS FILED – PART I</b>												
(Column 1)			(Column 2)		SMALL ENTITY <input type="checkbox"/>		OR			OTHER THAN SMALL ENTITY		
FOR		NUMBER FILED	NUMBER EXTRA		RATE (\$)	FEE (\$)	OR		RATE (\$)	FEE (\$)		
<input type="checkbox"/> BASIC FEE <small>(37 CFR 1.16(a), (b), or (c))</small>		N/A	N/A		N/A				N/A			
<input type="checkbox"/> SEARCH FEE <small>(37 CFR 1.16(k), (l), or (m))</small>		N/A	N/A		N/A		N/A					
<input type="checkbox"/> EXAMINATION FEE <small>(37 CFR 1.16(o), (p), or (q))</small>		N/A	N/A		N/A		N/A					
TOTAL CLAIMS <small>(37 CFR 1.16(i))</small>		minus 20 =	*		X \$ =		X \$ =					
INDEPENDENT CLAIMS <small>(37 CFR 1.16(h))</small>		minus 3 =	*		X \$ =		X \$ =					
<input type="checkbox"/> APPLICATION SIZE FEE <small>(37 CFR 1.16(s))</small>		If the specification and drawings exceed 100 sheets of paper, the application size fee due is \$250 (\$125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).										
<input type="checkbox"/> MULTIPLE DEPENDENT CLAIM PRESENT <small>(37 CFR 1.16(j))</small>												
* If the difference in column 1 is less than zero, enter "0" in column 2.												
<b>APPLICATION AS AMENDED – PART II</b>												
(Column 1)			(Column 2)		(Column 3)		SMALL ENTITY		OR		OTHER THAN SMALL ENTITY	
AMENDMENT	<b>12/22/2009</b>		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE (\$)	ADDITIONAL FEE (\$)		RATE (\$)	ADDITIONAL FEE (\$)	
	Total <small>(37 CFR 1.16(j))</small>		* 10	Minus	** 21	= 0	X \$ =		OR	X \$52=	0	
	Independent <small>(37 CFR 1.16(h))</small>		* 1	Minus	***3	= 0	X \$ =		OR	X \$220=	0	
	<input type="checkbox"/> Application Size Fee <small>(37 CFR 1.16(s))</small>											
	<input type="checkbox"/> FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM <small>(37 CFR 1.16(j))</small>											
TOTAL ADD'L FEE							OR		TOTAL ADD'L FEE			<b>0</b>
AMENDMENT			CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE (\$)	ADDITIONAL FEE (\$)		RATE (\$)	ADDITIONAL FEE (\$)	
	Total <small>(37 CFR 1.16(j))</small>		*	Minus	**	=	X \$ =		OR	X \$ =		
	Independent <small>(37 CFR 1.16(h))</small>		*	Minus	***	=	X \$ =		OR	X \$ =		
	<input type="checkbox"/> Application Size Fee <small>(37 CFR 1.16(s))</small>											
	<input type="checkbox"/> FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM <small>(37 CFR 1.16(j))</small>											
TOTAL ADD'L FEE							OR		TOTAL ADD'L FEE			
* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.												
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".												
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".												
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.												

Legal Instrument Examiner:  
/DIANIECE JACOBS/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.



UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents  
United States Patent and Trademark Office  
P.O. Box 1450  
Alexandria, VA 22313-1450  
www.uspto.gov

**COPY MAILED**

SEP 16 2009

**OFFICE OF PETITIONS**

VAN PELT, YI & JAMES, LLP  
10050 N. FOOTHILL BLVD #200  
CUPERTINO, CA 95014

In re Application of	:	
Ping Pan	:	
Application No. 11/354,569	:	DECISION ON PETITION
Filed: February 14, 2006	:	TO WITHDRAW FROM
Attorney Docket No. HAMMP008	:	RECORD

This is a decision on the Request to Withdraw as attorney or agent of record under 37 CFR § 1.36(b), filed July 24, 2009.

The request is **APPROVED**.

A grantable request to withdraw as attorney/agent of record must be signed by every attorney/agent seeking to withdraw or contain a clear indication that one attorney is signing on behalf of another/others.

The request was signed by Lee Van Pelt on behalf of all attorneys/agents of record who are associated with customer No. 21912. Therefore, Lee Van Pelt and all the attorneys/agents of record who are associated with customer No. 21912 have been withdrawn.

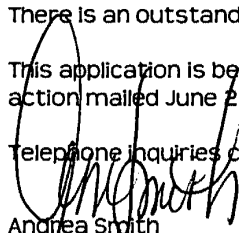
Applicant is reminded that there are no attorneys/agents of record at this time.

The request to change the correspondence of record is not acceptable<sup>1</sup> as the requested correspondence address is not that of: (1) the first named signing inventor; or (2) an intervening assignee of the entire interest under 37 C.F.R 3.71. All future communications from the Office will continue to be directed to the above listed address until otherwise properly notified by the applicant.

There is an outstanding Office action mailed June 22, 2009 that requires a reply from applicant.

This application is being referred to Technology Center Art Unit 2416 to await a response to the Office action mailed June 22, 2009.

Telephone inquiries concerning this decision should be directed to the undersigned at (571) 272-3226.

  
Andrea Smith  
Petitions Examiner  
Office of Petitions

cc: Greg Leitich  
804 Baylor Street  
Austin, TX 78703

<sup>1</sup> The Office will no longer change the correspondence address to that of a new practitioner unless the Request is accompanied by a power of attorney to a new practitioner (See USPTO Form PTO/SB/82).



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
11/354,569	02/14/2006	Ping Pan	HAMMP008

21912  
VAN PELT, YI & JAMES LLP  
10050 N. FOOTHILL BLVD #200  
CUPERTINO, CA 95014

**CONFIRMATION NO. 6912**  
**POWER OF ATTORNEY NOTICE**



Date Mailed: 09/15/2009

**NOTICE REGARDING CHANGE OF POWER OF ATTORNEY**

This is in response to the Power of Attorney filed 07/24/2009.

- The withdrawal as attorney in this application has been accepted. Future correspondence will be mailed to the new address of record. 37 CFR 1.33.

/amsmith/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.



**REQUEST FOR WITHDRAWAL  
AS ATTORNEY OR AGENT  
AND CHANGE OF  
CORRESPONDENCE ADDRESS**

Application Number	11/354,569
Filing Date	February 14, 2006
First Named Inventor	Ping Pan
Art Unit	4145
Examiner Name	Siming Liu
Attorney Docket Number	HAMMP008

To: Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Please withdraw me as attorney or agent for the above identified patent application, and

- all the practitioners of record;
- the practitioners (with registration numbers) of record listed on the attached paper(s); or
- the practitioners of record associated with Customer Number: 21912

**NOTE:** The immediately preceding box should only be marked when the practitioners were appointed using the listed Customer Number.

The reason(s) for this request are those described in 37 CFR :

- |   |  |  |  |
|---|--|--|--|
| <input type="checkbox"/> 10.40(b)(1)    | <input type="checkbox"/> 10.40(b)(2)     | <input type="checkbox"/> 10.40(b)(3)                       | <input type="checkbox"/> 10.40(b)(4)     |
| <input type="checkbox"/> 10.40(c)(1)(i) | <input type="checkbox"/> 10.40(c)(1)(ii) | <input type="checkbox"/> 10.40(c)(1)(iii)                  | <input type="checkbox"/> 10.40(c)(1)(iv) |
| <input type="checkbox"/> 10.40(c)(1)(v) | <input type="checkbox"/> 10.40(c)(1)(vi) | <input checked="" type="checkbox"/> 10.40(c)(2)            | <input type="checkbox"/> 10.40(c)(3)     |
| <input type="checkbox"/> 10.40(c)(4)    | <input type="checkbox"/> 10.40(c)(5)     | <input type="checkbox"/> 10.40(c)(6) Please explain below: |  |

**Certifications**

**Check each box below that is factually correct. WARNING: If a box is left unchecked, the request will likely not be approved.**

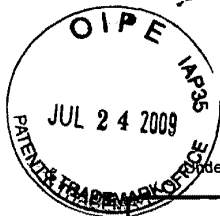
- I/We have given reasonable notice to the client, prior to the expiration of the response period, that the practitioner(s) intend to withdraw from employment.
- I/We have delivered to the client or a duly authorized representative of the client all papers and property (including funds) to which the client is entitled.
- I/We have notified the client of any responses that may be due and the time frame within which the client must respond.

Please provide an explanation, if necessary:

[Page 1 of 2]

This collection of information is required by 37 CFR 1.36. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.



Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

### REQUEST FOR WITHDRAWAL AS ATTORNEY OR AGENT AND CHANGE OF CORRESPONDENCE ADDRESS

Complete the following section only when the correspondence address will change. Changes of address will only be accepted to an inventor or an assignee that has properly made itself of record pursuant to 37 CFR 3.71.

Change the correspondence address and direct all future correspondence to:

A.  The address of the inventor or assignee associated with Customer Number: \_\_\_\_\_

OR

B.  Inventor or Assignee name Greg Leitch

Address 804 Baylor Street

City Austin State TX Zip 78703 Country US

Telephone \_\_\_\_\_ Email leitch@sbcglobal.net

I am authorized to sign on behalf of myself and all withdrawing practitioners.

Signature

Name Lee Van Pelt

Registration No. 38,352

Address 10050 N. Foothill Blvd., Suite 200

City Cupertino State CA Zip 95014 Country US

Date July 17, 2009 Telephone No. 408-973-2585

NOTE: Withdrawal is effective when approved rather than when received.

[Page 2 of 2]

This collection of information is required by 37 CFR 1.36. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	HAMMP008	6912
21912	7590	06/22/2009	EXAMINER	
VAN PELT, YI & JAMES LLP 10050 N. FOOTHILL BLVD #200 CUPERTINO, CA 95014			LIU, SIMING	
			ART UNIT	PAPER NUMBER
			2416	
			MAIL DATE	DELIVERY MODE
			06/22/2009	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b> 11/354,569	<b>Applicant(s)</b> PAN, PING	
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 2416	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1)  Responsive to communication(s) filed on 24 February 2009.
- 2a)  This action is **FINAL**.                      2b)  This action is non-final.
- 3)  Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4)  Claim(s) 1-21 is/are pending in the application.
  - 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5)  Claim(s) \_\_\_\_\_ is/are allowed.
- 6)  Claim(s) 1-21 is/are rejected.
- 7)  Claim(s) \_\_\_\_\_ is/are objected to.
- 8)  Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9)  The specification is objected to by the Examiner.
- 10)  The drawing(s) filed on \_\_\_\_\_ is/are: a)  accepted or b)  objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11)  The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12)  Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    - a)  All    b)  Some \*    c)  None of:
    - 1.  Certified copies of the priority documents have been received.
    - 2.  Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
    - 3.  Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1)  Notice of References Cited (PTO-892)
- 2)  Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3)  Information Disclosure Statement(s) (PTO/SB/08)  
 Paper No(s)/Mail Date \_\_\_\_\_.
- 4)  Interview Summary (PTO-413)  
 Paper No(s)/Mail Date. \_\_\_\_\_.
- 5)  Notice of Informal Patent Application
- 6)  Other: \_\_\_\_\_.

## DETAILED ACTION

This Action is in response to communication filled on 02/24/2009.

### ***Claim Rejections - 35 USC § 101***

Base on the amendment, 101 rejections are removed to claims 17-21.

### ***Response to Arguments***

1. Applicant's arguments with respect to claims 1, 11, 17 have been considered but are moot in view of the new ground(s) of rejection. Applicant amended the claims 1, 11, 17 to include new limitation "the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire", which necessitates the new ground of rejection.

### ***Claim Rejections - 35 USC § 103***

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.



2. Claims 1-4, 7, 11-12, 15, 17, 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang US 2003/0117950 A1, in view of Voit US 2006/0047851 A1, further in view of Blanchet US 2004/0133692 A1, further in view of Sridhar US 2006/0018252 A1.

3. Regarding claim 1, Huang teaches a method of providing protection to network traffic (Huang, page 2, [0015], lines 1-6), comprising:  
sending (Huang, page 2, [0016], right column, line 1: "receiving a request to set up". There must be sending, thus receiving can happen) a ... protection configuration parameter (Huang, page 2, [0016], right column, lines 2-4: "the request specifying a required protection bandwidth for the label switched path segment", "required protection bandwidth" can be considered as a protection configuration parameter) for configuring a standby ... between a source node and a destination node (Huang, page 2, [0016], right column, lines 4-5: "and determining a backup route to the tail end node"), ...; receiving a ... configuration acknowledgement indicating whether the ... protection configuration parameter has been accepted by the destination node; and  
in the event that the ... protection configuration parameter has been accepted by the destination node, using the standby ... (Huang, page 2, [0016], right column, lines 6-14: In response of the request of setting up a label switched path segment over a direct connection between two nodes, a backup route is also being determined);  
wherein the standby Pseudowire is configured based at least in part on the ... protection configuration parameter parameter (Huang, page 2, [0016], right column, lines 8-12: "The method also includes signaling to reserve the required protection bandwidth along

the backup route, receiving confirmation of reservation of the required protection bandwidth and generating a backup connection map", required protection bandwidth as a configuration parameter is a major factor in the backup path forming).

Huang doesn't expressly teach Pseudowire and Pseudowire protection.

Voit teaches Pseudowire (Voit, page 2, [0011], lines 2-7) and Pseudowire protection (Voit, page 4, [0046], lines 1-3: "a network topology is provided with redundant pseudowire connections ...").

At the time of the invention was made, it would have been obvious to a person of ordinary skill in the art to implement Pseudowire as a type of network service in the system disclosed by Huang in order to increase communication security since Voit teaches that PWs can provide point-to-point connectivity and are similar to virtual private link. Voit also teaches providing data traffic protection for primary Pseudowire path (Voit, page 4, [0046], lines 1-3). Both Huang and Voit are in the same field of endeavor (network transfer) and are directed to the same problem sought to be solved (data traffic protection).

Huang in view of Voit doesn't expressly teach that receiving a configuration acknowledgement indicating whether the configuration parameter has been accepted by the destination node.

Blanchet teaches that receiving a configuration acknowledgement indicating whether the configuration parameter has been accepted by the destination node (Blanchet, page 4, [0035], lines 2-4).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the system to send an ACK indicating the acceptance of the configuration parameters in the system disclosed by Huang in view of Voit in order to make the system more reliable. Both Huang in view of Voit and Blanchet are in the same field of endeavor (Network transfer).

Huang in view of Voit and Blanchet doesn't expressly teach that the configuration parameter indicating a protection property associated with the standby link.

Sridhar teaches that the configuration parameter indicating a protection property associated with the standby link (Sridhar, [0031], lines 3-10: it is noted the backup path setup is based on one of the three protection schemes, the parameter indicate the type of protection scheme is considered as the configuration parameters).

At the time of the invention was made, it would have been obvious to a person of ordinary skill in the art to configure a backup path based on the configuration parameters which indicate a property associated with the standby link in the system disclosed by Huang in view of Voit and Blanchet in order increase the efficiency of the system. Since not all primary paths have the same risk, it enables the system to weight the risk and assign proper resources to each primary path for protection. Both Huang in view of Voit, Blanchet and Sridhar are in the same field of endeavor (Network transfer) and are directed to the same problem sought to be solved (data traffic protection).

2. Regarding claims 2, 12, Huang in view of Voit, Blanchet and Sridhar further teaches the standby (Huang, page 2, [0016], right column, lines 5: "backup" is

equivalent to standby in the context) Pseudowire (VOIT, [0002], lines 1-2) is configured to provide protection to at least one primary (Huang, page 1, [0008], lines 2-4) Pseudowire.

3. Regarding claim 3, Huang in view of Voit, Blanchet and Sridhar further teaches the standby Pseudowire is configured to provide protection to at least one primary Pseudowire (Huang, page 1, [0008], lines 2-4), and in the event that the primary Pseudowire (VOIT, [0002], lines 1-2) fails to transfer network traffic (Huang, page 2, [0010], line 5: "when a fault is discovered in a single link between two nodes along a path"), switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire (Huang, page 2, [0010], lines 9-10: "switches the traffic that was using the connection to the alternate path", the limitation "Pseudowire" has been discussed).

4. Regarding claim 4, Huang in view of Voit, Blanchet and Sridhar further teaches the standby Pseudowire is dynamically selected from a plurality of connections (Huang, page 4, [0040], lines 12-14: "The backup route may, for instance, be selected from a table of routes that have been pre-computed to connect the head end node 102A to the tail end node 102B"; page 4, [0040], right column, lines 1-2: "a backup route can be determined instantaneously by the head end node 102A given information about the current state of the network 100").

5. Regarding claims 7, 15, 20, Huang in view of Voit, Blanchet and Sridhar further teaches the Pseudowire protection configuration parameter includes a protection scheme (Sridar, [0031], lines 3-10: it is noted the backup path setup is based on one of the three protection schemes, the parameter indicate the type of protection scheme is considered as the configuration parameters).

6. Regarding claim 10, Huang in view of Voit, Blanchet and Sridhar further teaches the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP) (Huang, page 1, [0005], right column, last 3 lines).

7. Regarding claim 11, Huang in view of Voit, Blanchet and Sridhar teaches a system for providing protection to network traffic, comprising:  
a processor (Blanchet, Fig. 2, element 50: It's inherent, since all computers have at least one processor) configured to:  
send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire; and  
receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, use the standby Pseudowire;

wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter; and a memory coupled to the processor, configured to provide the processor with instructions (Blanchet, Fig. 2, element 50: it's inherent since all computers have a memory coupled to a processor, and provide instructions with processor). (All of the remaining limitations have been discussed in claim 1)

8. Regarding claim 17, Huang in view of Voit, Blanchet and Sridhar teaches a computer program product (Huang, page 2, [0013]: "the 'gold' level of service protection" is a computer program product) for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions (Blanchet, Fig. 2, element 50: it's inherent, since all computers have memory and can give computer instructions) for:

sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire;

receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire; wherein the standby

Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter (All of the remaining limitations have been discussed in claim 1).

9. Claims 5, 13, 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over under 35 U.S.C. 103(a) as being unpatentable over Huang, in view of Voit, Blanchet and Sridhar, further in view of Cruz, US 2006/0046658 A1.

10. Regarding claims 5, 13, 18, Huang in view of VOIT, Blanchet and Sridhar as applied in claim above teaches a method as recited in Claim 1, wherein the Pseudowire (VOIT, [0002], lines 1-2) protection configuration parameter (Huang, page 2, [0016], right column, lines 2-4: “the request specifying a required protection bandwidth for the label switched path segment”, “required protection bandwidth” is equivalent to a protection configuration parameter) includes ...

Huang in view of Voit, Blanchet and Sridhar doesn't expressly teach that a domain type.

Cruz teaches a domain type (Cruz, page 1, para 0017, line 2: According to the specification of the application, domain type is about whether the network is either multi-hop or single hop).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include domain type. The reason is that by including domain type in the configuration parameter, it would be more accurate to select a desire standby path, given that you have more information about

the network. The method of change the configuration parameter by including the domain type of Huang in view of VOIT, Blanchet and Sridhar was within the ordinary ability of one of ordinary skill in the art based on the teachings of Cruz.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, VOIT, Blanchet, Sridhar and Cruz to obtain the invention as specified in claims 5, 13, 18.

11. Claims 6, 14, 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang in view of VOIT, Blanchet and Sridhar, further in view of Rathunde, US 6,574,477 B1.

12. Regarding claims 6, 14, 19, Huang in view of Voit, Blanchet and Sridhar teaches a method as recited in Claim 1, wherein the Pseudowire protection configuration parameter (previous discussed) includes ...

Huang in view of Voit, Blanchet and Sridhar doesn't expressly teach that a protection type.

Rathunde teaches a protection type (Rathunde, col 9, line 3: "type of standby mode", according to the specification of the application, protection type just means what type of standby mode).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include a protection type. The reason is that by including protection type in the configuration parameter, it would be more accurate to select a desire standby path, given that you have more information



about the network. The method of change the configuration parameter by including the protection type of Huang in view of VOIT and Blanchet was within the ordinary ability of one of ordinary skill in the art based on the teachings of Rathunde.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, VOIT, Blanchet and Rathunde to obtain the invention as specified in claims 6, 14, 19.

13. Claims 8-9, 16, 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang, Voit, Blanchet and Sridhar, further in view of Saleh, US 7,200,104 B2.

14. Regarding claims 8, 16, 21, Huang in view of Voit, Blanchet and Sridhar teaches a method as recited in Claim 1, wherein the Pseudowire protection configuration parameter (previous discussed) includes a ...

Huang in view of Voit, Blanchet and Sridhar doesn't expressly teach that a priority.

Saleh teaches a priority (Saleh, col 3, line 38: "restoration priority level").

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include priority. The reason is that by including domain type in the configuration parameter, it would be more accurate to select a desire standby path, given that you have more information about the network. The method of change the configuration parameter by including the domain type of Huang in view of Voit and Blanchet was within the ordinary ability of one of ordinary skill in the art based on the teachings of Saleh.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, Voit, Blanchet and Saleh to obtain the invention as specified in claims 8, 16, 21.

15. Regarding claim 9, Huang in view of Voit, Blanchet and Sridhar teaches a method as recited in Claim 1, further including determining whether to preempt existing traffic on the standby Pseudowire, the determination being based at least in part on a priority associated with the standby Pseudowire (Saleh, col 3, lines 3-8: QoS is equivalent to priority).

### ***Conclusion***

16. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

17. Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIMING LIU whose telephone number is (571)270-3859. The examiner can normally be reached on Monday-Friday 8:30am-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, William Trost can be reached on 571-272-7872. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/S. L./  
Examiner, Art Unit 2416

Application/Control Number: 11/354,569  
Art Unit: 2416

Page 14

/William Trost/  
Supervisory Patent Examiner, Art Unit 2416

<b>Notice of References Cited</b>	Application/Control No. 11/354,569	Applicant(s)/Patent Under Reexamination PAN, PING	
	Examiner SIMING LIU	Art Unit 2416	Page 1 of 1

**U.S. PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A US-2006/0018252	01-2006	Sridhar et al.	370/216
*	B US-2006/0047851	03-2006	Voit et al.	709/239
C	US-			
D	US-			
E	US-			
F	US-			
G	US-			
H	US-			
I	US-			
J	US-			
K	US-			
L	US-			
M	US-			

**FOREIGN PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
N					
O					
P					
Q					
R					
S					
T					

**NON-PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)				
U					
V					
W					
X					


\*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)  
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

<b>Index of Claims</b>  	<b>Application/Control No.</b> 11354569	<b>Applicant(s)/Patent Under Reexamination</b> PAN, PING
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 2416

✓	<b>Rejected</b>	-	<b>Cancelled</b>	N	<b>Non-Elected</b>	A	<b>Appeal</b>
=	<b>Allowed</b>	÷	<b>Restricted</b>	I	<b>Interference</b>	O	<b>Objected</b>

Claims renumbered in the same order as presented by applicant
  CPA
  T.D.
  R.1.47

CLAIM		DATE							
Final	Original	10/30/2008	06/17/2009						
	1	✓	✓						
	2	✓	✓						
	3	✓	✓						
	4	✓	✓						
	5	✓	✓						
	6	✓	✓						
	7	✓	✓						
	8	✓	✓						
	9	✓	✓						
	10	✓	✓						
	11	✓	✓						
	12	✓	✓						
	13	✓	✓						
	14	✓	✓						
	15	✓	✓						
	16	✓	✓						
	17	✓	✓						
	18	✓	✓						
	19	✓	✓						
	20	✓	✓						
	21	✓	✓						

<b>Search Notes</b>  	<b>Application/Control No.</b>  11354569	<b>Applicant(s)/Patent Under Reexamination</b>  PAN, PING
	<b>Examiner</b>  SIMING LIU	<b>Art Unit</b>  2416

SEARCHED			
Class	Subclass	Date	Examiner
370	216, 225, 228	10/30/2008	/SL/
709	220	10/30/2008	/SL/
above	update search	6/17/2009	/SL/

SEARCH NOTES		
Search Notes	Date	Examiner
East Class search	11/10/2008 update 6/17/2009	/SL/
Palm inventor name search	10/30/2008 update 6/17/2009	/SL/
Consulted 101 issues with Peng, John	11/10/2008	/SL/

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner

--	--

## EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L4	2421	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
L5	3849	(709/220).ccls.	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
L6	291	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
L7	5	(709/220).ccls. and L6	US-PGPUB; USPAT	OR	ON	2009/06/17 14:35
S5	115	pseudowire	US-PGPUB; USPAT	OR	ON	2008/10/08 12:53
S6	0	pseudowire and tele	US-PGPUB; USPAT	OR	ON	2008/10/08 12:53
S7	217	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2008/10/08 13:07
S8	9	S7 with protection	US-PGPUB; USPAT	OR	ON	2008/10/08 13:08
S9	4	S7 with protection and @ad- "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 13:09
S10	1	"20040223498".pn.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:07
S11	0	(pseudowire or pseudo-wire) and initiliz\$5	US-PGPUB; USPAT	OR	ON	2008/10/08 14:14
S12	133	(pseudowire or pseudo-wire) and initi\$5	US-PGPUB; USPAT	OR	ON	2008/10/08 14:15
S13	51	(pseudowire or pseudo-wire) and initi\$5 and @ad- "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:15
S14	2193	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:17
S15	6	(370/216,225,228).ccls. and S7	US-PGPUB; USPAT	OR	ON	2008/10/08 14:23
S16	3	(709/220).ccls. and S7	US-PGPUB; USPAT	OR	ON	2008/10/08 14:26
S17	31	((PING) near2 (PAN)).INV.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:32
S18	2	((PING) near2 (PAN)).INV. and pseudowire	US-PGPUB; USPAT	OR	ON	2008/10/08 14:33
S19	2	((PING) near2 (PAN)).INV. and (pseudowire).clm.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:33
S20	66	S7 and (primary)	US-PGPUB; USPAT	OR	ON	2008/10/08 14:38
S21	23	S7 and (primary) and @ad- "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:39
S22	75	S7 and (config\$7) and @ad- "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:44



S23	2	TDM pseudowire	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:09
S24	106	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol))	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:16
S26	43	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:17
S27	11	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (standby or backup)	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:19
S28	9	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (primary or main) and (secondly or backup or standby)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 09:00
S29	43	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:34
S30	14	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:35
S31	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:38
S32	18	(pseudowire or (pseudo wire) or pseudo-wire) and ((config\$7) same (destination near5 node))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:38
S33	1	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with acknowledgement)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:41
S34	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:44
S35	8	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:44
S36	370	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:47

S37	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S33) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:48
S38	218	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2008/10/09 10:49
S39	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S38) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:49
S40	275	pseudowire or pseudo-wire or (pseudo wire)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:49
S41	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S40) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:50
S42	233	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:50
S43	27	S40 and initialization	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:54
S44	3434011	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2008/10/29 09:26
S46	1334019	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:27
S47	3640734	(alter\$7 or backup or standby)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:28
S48	29788561	@ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:28
S49	6386553	(pick\$5 or select\$5 or choos\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:30
S50	409	(S44 near7 S46) with (S49 near7 S47 near7 S44) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:31
S51	238	(S44 near7 S46) with (S49 near7 S47 near7 S44) and S48 and (priority or bandwidth)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:38

S52	25	(S44 near7 S46) with (S49 near7 S47 near7 S44) same (priority or bandwidth or parameter) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:43
S53	2289	S47 with config\$7 with (primary near7 S47)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:49
S54	159	(S47 near5 S44) with config \$7 with (primary near7 S47)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:54
S55	175	(S47 near5 S44) with config \$7 with (primary near7 S44)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:54
S56	111	(S47 near5 S44) with config \$7 with (primary near7 S44) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:56
S57	7	09/859166	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 10:26
S58	33	(restoration scheme) and ("1:N")	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 10:50
S59	4	(restoration scheme) and (priority) and (standby mode)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 12:42
S60	18	(restoration scheme) and (priority) and (config\$7 near5 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:27
S61	3706723	(send\$7 or transmit\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:30
S62	0	(source node) with S61 with (config\$7 near3 parameter \$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:31
S63	5	(source) with S61 with (config\$7 near3 parameter \$1) with (destin\$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:32
S64	569	(source) with S61 with (parameter\$1) with (destin \$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:33

S65	54	(source node) with S61 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:33
S66	2959	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:53
S67	0	(ack or acknowledgement) same (config\$7 parameter\$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:53
S68	20807	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:54
S69	52906	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:54
S70	29	(ack or acknowledgement) and (restoration scheme) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:55
S71	137	S61 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:19
S72	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:29
S73	10549	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:29
S74	759	handshaking and @ad<"20050214" and (S44 with S46)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:30
S75	2	"6553034".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:32
S76	108	(virtual path) and ((protection or restoration) near5 scheme) and priority	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:34
S77	103	(virtual path) and ((protection or restoration) near5 scheme) and priority and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:34

S78	3479	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:51
S79	2628	(protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:52
S80	6	((protection or restoration) near5 parameter) with (S61) and (destin\$7 near3 node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:53
S81	26	((protection or restoration) near5 parameter) and (handshaking) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:55
S82	73	((protection or restoration) near5 parameter) and (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:04
S83	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:04
S84	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S85	15	((protection or restoration or config\$7) near5 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S86	4	((protection or restoration or config\$7) near2 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S87	651	handshaking and @ad<"20050214" and (config\$7 parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:10
S88	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:27
S89	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:27
S90	276	receiv\$7 acknowledgement destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:28

S91	0	receiv\$7 acknowledgement (parameter accept\$5 destination node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:40
S92	5	receiv\$7 acknowledgement accept\$3 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:45
S93	7	receiv\$7 acknowledgement parameter accept\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:48
S94	2	"20030117950".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:20
S95	771	(domain type) with (parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:25
S96	3	(parameter) near5 includ\$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:26
S97	3	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:27
S98	64	(parameter\$1) with (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:27
S99	0	(domain type) with (single- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:28
S100	0	(domain type) with (single near5 hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:29
S101	64	(domain type) with parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:29
S102	179	(single-hop) same (multi- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:32
S103	9	(single-hop) same (multi- hop) same (parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:33

S104	0	field with indica\$7 with ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:34
S105	147	field with indica\$7 with (topology)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:34
S106	6	field with indica\$7 with (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:36
S107	179	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:37
S108	10	(field or parameter) same ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:37
S109	283	((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:40
S111	21	(parameter or field) same S109	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:41
S112	0	S109 same (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:44
S113	134	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:46
S114	0	parameter with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:48
S115	0	(field or parameter) with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:49
S116	0	(field or parameter) with (show\$3 or indicat\$5) same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:49

S117	0	(field or parameter) with (domain type) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S118	68	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S119	14	(field or parameter) with (indicat\$5 or show\$5) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S120	1	(protection type) and (standby path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:44
S121	2636	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:46
S122	283	(hot and cold) same standby and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:47
S123	51	(hot and cold) and (parameter with standby) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:48
S124	20	(field with indicat\$5 with (standby mode)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:49
S126	696	config\$9 with (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:50
S127	406	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:51
S128	324	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:51
S129	194	config\$9 near3 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:52
S130	7	config\$9 near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:54



S131	9	type with (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:01
S132	4	type near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:01
S133	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:38
S135	134	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:39
S136	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:43
S137	126	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:43
S138	11	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) same (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:44
S139	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 netowrk)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:53
S140	136	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 network)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:53
S141	0	(paramete or field or bit) with indicat\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:55
S142	2	(paramete or field or bit) with indicat\$7 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:55
S143	1	(protection near3 properties) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2009/05/02 14:27
S144	7	(protection near3 (parameter or propert\$5)) and standby near3 path	US-PGPUB; USPAT; EPO; JPO	OR	ON	2009/05/02 14:28

S145	61963	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	OR	OFF	2009/05/02 15:12
S146	15	configu\$5 with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	OFF	2009/05/02 15:13
S147	25	(pseudowire or pseudo-wire) and (standby)	US-PGPUB; USPAT	OR	ON	2009/05/02 15:20
S148	3	(pseudowire or pseudo-wire) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 15:20
S149	12	(protection scheme) and (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:04
S150	1	(protection scheme) with (standby path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:06
S151	1	(protection (type or property)) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:22
S152	10	((protection (type or property)) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:23
S153	19	( (type or property) or QOS) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:26
S154	1	(protection scheme) with (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:27
S155	2	(protection scheme) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:27
S156	3	(protection (scheme or propert\$3 or parameter or type)) same (standby (path or route))	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/05/02 16:29
S157	20	(backup path) with (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:22
S158	0	(backup path) with (protection near3 parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:35
S159	121	(backup path) with (protection )	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:35
S160	1	(genera\$5 or configur\$5) with (backup path) with (parameter)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:36
S161	144	(genera\$5 or configur\$5) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:37

S162	22	(genera\$5 or configur\$5) with (backup path) and (protection scheme)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:38
S163	0	(genera\$5 or configur\$5) with (backup path) with (base or according)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:42
S164	0	(genera\$5 or configur\$5) with (backup path) with ("base" or "according")	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:42
S165	159	(genera\$5 or configur\$5 or setup) with (backup path)	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:43
S166	15	(genera\$5 or configur\$5 or setup) with (backup path) not S161	US-PGPUB; USPAT; EPO; JPO	ADJ	ON	2009/06/17 10:43

6/ 17/ 2009 2:36:23 PM

C:\ Documents and Settings\ sliu3\ My Documents\ EAST\ Workspaces\ 11354569.wsp

JFW



PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor:	Ping Pan	Examiner:	Siming Liu
Application No.:	11/354,569	Art Unit:	4145
Filed:	February 14, 2006	Docket No.	HAMMP008
Title:	PSEUDOWIRE PROTECTION		

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in a prepaid envelope addressed to: Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on:

2/19, 2009.

Elaine Nguyen

TRANSMITTAL OF AMENDMENT A

Mail Stop Amendment  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

Transmitted herewith is Amendment A in response to Office Action mailed November 20, 2008 in the above-identified application.

The fee has been calculated as shown below.

CLAIMS	After Amd.	HP*	Extra	Small Entity		OR	Large Entity	
				Rate	Fee		Rate	Fee
Total	21	21	-0-	x \$26 = \$		OR	x \$52 = \$	
Independent	3	3	-0-	x \$110 = \$		OR	x \$220 = \$	
Multiple Dependent Claims				x \$195 = \$		OR	x \$390 = \$	
*HP = Highest previously paid				TOTAL FEE \$		OR	TOTAL FEE \$	-0-

Applicant(s) hereby petition for following extension of time in which to respond to the outstanding Office Action.

	SMALL ENTITY		OR	LARGE ENTITY	
	Rate	Add'l Fee		Rate	Add'l Fee
<input type="checkbox"/> Extension for Response within FIRST month	x \$65 = \$		OR	x \$130 = \$	
<input type="checkbox"/> Extension for Response within SECOND month	x \$245 = \$		OR	x \$490 = \$	
<input type="checkbox"/> Extension for Response within THIRD month	x \$555 = \$		OR	x \$1110 = \$	
<input type="checkbox"/> Extension for Response within FOURTH month	x \$865 = \$		OR	x \$1730 = \$	
<input type="checkbox"/> Extension for Response within FIFTH month	x \$1175 = \$		OR	x \$2350 = \$	

Applicant(s) believe that no (additional) Extension of Time is required; however, if it is determined that such an extension is required, Applicant(s) hereby petition that such an extension be granted and authorize the Commissioner to charge the required fees for an Extension of Time under 37 CFR 1.136 to Deposit Account No. 50-0685. (HAMMP008 ).

Enclosed is our Check No. \_\_\_ in the amount of \$ \_\_\_\_\_ to cover the additional claim fee and/or extension of time fees.

Enclosed is Applicant Initiated Interview Request Form, PTOL-413A.

Enclosed are \_\_\_\_\_ sheets replacement drawings.

Please charge Deposit Account No. 50-0685 (HAMMP008 ) in the amount of \$ \_\_\_\_\_ to cover the additional claim fee and/or extension of time fees.

If the required fees are missing or any additional fees are required during the pendency of the subject application, please charge such fees or credit any overpayment to Deposit Account No. 50-0685 (HAMMP008 ).

OTHER:

Respectfully submitted,  
VAN PELT, YI & JAMES LLP



Diana Y. Fu  
Registration No. 52,924  
V 408-973-2593  
F 408-973-2595

10050 N. Foothill Blvd., Suite 200  
Cupertino, CA 95014




IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor:	Ping Pan	Examiner:	Siming Liu
Application No.:	11/354,569	Art Unit:	4145
Filed:	February 14, 2006	Docket No.:	HAMMP008
Title:	PSEUDOWIRE PROTECTION		

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in a prepaid envelope addressed to: Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on:

2/19, 2009.

  
Elaine Nguyen

**AMENDMENT A**

Mail Stop Amendment  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Dear Sir:

This is in response to the Office Action mailed November 20, 2008. The following amendments and remarks are respectfully submitted.

JFW



PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor:	Ping Pan	Examiner:	Siming Liu
Application No.:	11/354,569	Art Unit:	4145
Filed:	February 14, 2006	Docket No.	HAMMP008
Title:	PSEUDOWIRE PROTECTION		

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in a prepaid envelope addressed to: Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on:

2/19, 2009.

Elaine Nguyen

TRANSMITTAL OF AMENDMENT A

Mail Stop Amendment  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

Transmitted herewith is Amendment A in response to Office Action mailed November 20, 2008 in the above-identified application.

The fee has been calculated as shown below.

CLAIMS	After Amd.	HP*	Extra	Small Entity		Large Entity	
				Rate	Fee	Rate	Fee
Total	21	21	-0-	x \$26 = \$		OR	x \$52 = \$
Independent	3	3	-0-	x \$110 = \$		OR	x \$220 = \$
Multiple Dependent Claims				x \$195 = \$		OR	x \$390 = \$
*HP = Highest previously paid				TOTAL FEE \$		OR	TOTAL FEE \$ -0-

Applicant(s) hereby petition for following extension of time in which to respond to the outstanding Office Action.

	SMALL ENTITY		OR	LARGE ENTITY	
	Rate	Add'l Fee		Rate	Add'l Fee
<input type="checkbox"/> Extension for Response within FIRST month	x \$65 = \$		OR	x \$130 = \$	
<input type="checkbox"/> Extension for Response within SECOND month	x \$245 = \$		OR	x \$490 = \$	
<input type="checkbox"/> Extension for Response within THIRD month	x \$555 = \$		OR	x \$1110 = \$	
<input type="checkbox"/> Extension for Response within FOURTH month	x \$865 = \$		OR	x \$1730 = \$	
<input type="checkbox"/> Extension for Response within FIFTH month	x \$1175 = \$		OR	x \$2350 = \$	

Applicant(s) believe that no (additional) Extension of Time is required; however, if it is determined that such an extension is required, Applicant(s) hereby petition that such an extension be granted and authorize the Commissioner to charge the required fees for an Extension of Time under 37 CFR 1.136 to Deposit Account No. 50-0685. (HAMMP008 ).

Enclosed is our Check No. \_\_\_ in the amount of \$ \_\_\_\_\_ to cover the additional claim fee and/or extension of time fees.

Enclosed is Applicant Initiated Interview Request Form, PTOL-413A.

Enclosed are \_\_\_\_\_ sheets replacement drawings.

Please charge Deposit Account No. 50-0685 (HAMMP008 ) in the amount of \$ \_\_\_\_\_ to cover the additional claim fee and/or extension of time fees.

If the required fees are missing or any additional fees are required during the pendency of the subject application, please charge such fees or credit any overpayment to Deposit Account No. 50-0685 (HAMMP008 ).

OTHER:

Respectfully submitted,  
VAN PELT, YI & JAMES LLP



Diana Y. Fu  
Registration No. 52,924  
V 408-973-2593  
F 408-973-2595

10050 N. Foothill Blvd., Suite 200  
Cupertino, CA 95014



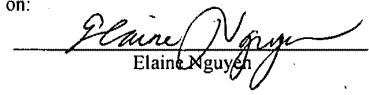


IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor:	Ping Pan	Examiner:	Siming Liu
Application No.:	11/354,569	Art Unit:	4145
Filed:	February 14, 2006	Docket No.:	HAMMP008
Title:	PSEUDOWIRE PROTECTION		

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in a prepaid envelope addressed to: Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on:

2/19, 2009.   
Elaine Nguyen

**AMENDMENT A**

Mail Stop Amendment  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Dear Sir:

This is in response to the Office Action mailed November 20, 2008. The following amendments and remarks are respectfully submitted.

## **AMENDMENTS TO THE ABSTRACT**

Please replace the section entitled "Abstract of the Disclosure" beginning on page 17 with the following replacement section:

### **PSEUDOWIRE PROTECTION**

#### **ABSTRACT OF THE DISCLOSURE**

Providing protection to network traffic includes sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node, and in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire, wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter.

**AMENDMENTS TO THE SPECIFICATION**

Please replace the title of the invention appearing on the Cover Page and Page 1 with the following:

~~PSEUDOWIRE PROTECTION~~

PSEUDOWIRE PROTECTION USING A STANDBY PSEUDOWIRE

## IN THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application:

### Listing of Claims:

1. (Currently Amended) A method of providing protection to network traffic, comprising:  
    sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire;  
    receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and  
    in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire;  
    wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter.
2. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.
3. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.
4. (Original) A method as recited in Claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.
5. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a domain type.
6. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a protection type.
7. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a protection scheme.
8. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a priority.

9. (Original) A method as recited in Claim 1, further including determining whether to preempt existing traffic on the standby Pseudowire, the determination being based at least in part on a priority associated with the standby Pseudowire.
10. (Original) A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP).
11. (Currently Amended) A system for providing protection to network traffic, comprising:  
a processor configured to:  
send a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire; and  
receive a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and  
in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, use the standby Pseudowire;  
wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter; and  
a memory coupled to the processor, configured to provide the processor with instructions.
12. (Original) A system as recited in Claim 11, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.
13. (Original) A system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a domain type.
14. (Original) A system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a protection type.
15. (Original) A system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a protection scheme.
16. (Original) A system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a priority.

17. (Currently Amended) A computer program product for configuring a Pseudowire between a source node and a destination node, the computer program product being embodied in a computer readable storage medium and comprising computer instructions for:

    sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, the Pseudowire protection configuration parameter indicating a protection property associated with the standby Pseudowire;

    receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and

    in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire;

    wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter.

18. (Original) A computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a domain type.

19. (Original) A computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a protection type.

20. (Original) A computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a protection scheme.

21. (Original) A computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a priority.

## REMARKS

Claims 1, 11, and 17 have been amended to clarify the subject matter regarded as the invention. Claims 1-20 are pending.

### *Claim Rejections – 35 U.S.C. §101*

Claims 17-21 stand rejected under 35 U.S.C. §101 for being directed to non-statutory subject matter. The amended claims are believed to overcome the rejection.

### *Claim Rejections – 35 U.S.C. §103*

Claims 1-4, 7, 11-12, 15, 17, and 20 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Huang (US 2003/0117950) and in view of Background of the Inventions of the present application and further in view of Blanchet (US 2004/0133692).

With respect to Claims 1, 11, and 17, support for the amendment may be found, as an example and without limitation, at paragraph [0021] of the present application.

Neither Huang, nor Blanchet, nor the Background of the Invention teaches, either singly or in combination, that the Pseudowire protection configuration parameter indicates a protection property associated with the standby Pseudowire. Claims 1, 11, and 17 are therefore believed to be allowable.

Further, The Office Action referred to Paragraph [0002], lines 3-4 and lines 8-10 of the present application as admission of prior art (AAPA), stating that “at the time of the invention, it would have been obvious to a person of ordinary skill in the art to implement Pseudowire as a type of network service. The reason is that Pseudowire can emulate the operation of a “transparent wire” carrying the native service. The method of modifying the system of Huang was within the ordinary ability of one of ordinary skill in the art based on the teaching of AAPA.” Applicant respectfully disagrees with this line of reasoning.

“Under § 103, the scope and content of the prior art are to be determined; differences between the prior art and the claims at issue are to be ascertained; and the level of ordinary skill in the pertinent art resolved. Against this background the obviousness or nonobviousness of the

subject matter is determined. Such secondary considerations as commercial success, long felt but unsolved needs, failure of others, etc., might be utilized to give light to the circumstances surrounding the origin of the subject matter sought to be patented.” *Graham v. John Deere Co.*, 383 U.S. at 17-18, (1966).

The Background of the Invention explicitly describes such long felt but unresolved needs, and failure of others to provide adequate Pseudowire protection:

**[0003]** At the edge of a network, a network edge device such as an edge router may receive multiple Layer-2 flows (also referred to as Attachment Circuits (ACs)). In a typical network supporting Pseudowires, each AC is mapped to a Pseudowire. Ingress packets received mapped to a specific Pseudowire are labeled with an identifier associated with this Pseudowire, and are switched via the Pseudowire. A physical link may support one or more Pseudowires. Ideally, the data flow in a Pseudowire should be protected. In other words, if an active Pseudowire fails, the data flow should be redirected to an alternative Pseudowire to avoid data loss.

**[0004]** Pseudowires can operate over many physical media types. However, existing Pseudowire systems typically provide no protection or very limited protection. For example, there is usually no data protection for Pseudowires on different physical media types, since most network protection schemes, such as APS for SONET, Link Aggregation for Ethernet, do not apply over multiple physical media types.

**[0005]** Some MPLS devices implement schemes such as MPLS Fast Reroute to provide limited data protection. These existing schemes, however, often do not provide adequate protection. Take the following scenario as an example: between two provider edges (PEs), a first tunnel comprising multiple Pseudowires is protected by a second tunnel. Due to network topology constraints, the two tunnels may have different bandwidth. This is a possible scenario in an MPLS Fast Reroute operation. In this example, the second tunnel may have lower bandwidth than that of the first one. If the first tunnel should fail, the amount of data that needs to be redirected through the second tunnel may exceed the capacity of the second tunnel. Furthermore, existing protocols typically do not provide a way of determining which data gets priority. Thus, certain mission critical data may be dropped while other less critical data may pass through.

**[0006]** It would be desirable to have a way to provide better Pseudowire protection and to have more control during switchover. It would also be desirable if the protection scheme could be implemented without significant changes to existing protocols and devices.



Thus, the Background of the Invention shows that at the time of the invention, given the long felt but unresolved needs and failure of others, it would not have been obvious for one with ordinary skill in the art to combine a Pseudowire as taught in the Background with a system such as that of Huang.

As such, Claims 1, 11, and 17 are believed to be allowable of Huang, the Background of the Invention, and Blanchet.

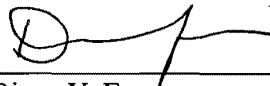
Claims 2-10 depend from Claim 1, Claims 12-16 depend from Claim 11, and Claims 17-21 depend from Claim 17. They are believed to be allowable for the same reasons described above.

The foregoing amendments are not to be taken as an admission of unpatentability of any of the claims prior to the amendments.

Reconsideration of the application and allowance of all claims are respectfully requested based on the preceding remarks. If at any time the Examiner believes that an interview would be helpful, please contact the undersigned.

Dated: 2/19/09

Respectfully submitted,



Diana Y. Fu  
Registration No. 52,924  
V 408-973-2593  
F 408-973-2595

VAN PELT, YI & JAMES LLP  
10050 N. Foothill Blvd., Suite 200  
Cupertino, CA 95014

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

<b>PATENT APPLICATION FEE DETERMINATION RECORD</b> Substitute for Form PTO-875					Application or Docket Number <b>11/354,569</b>		Filing Date <b>02/14/2006</b>		<input type="checkbox"/> To be Mailed			
<b>APPLICATION AS FILED – PART I</b>												
(Column 1)			(Column 2)		SMALL ENTITY <input type="checkbox"/>		OR			OTHER THAN SMALL ENTITY		
FOR		NUMBER FILED	NUMBER EXTRA		RATE (\$)	FEE (\$)	OR	RATE (\$)	FEE (\$)			
<input type="checkbox"/> BASIC FEE <small>(37 CFR 1.16(a), (b), or (c))</small>		N/A	N/A		N/A			N/A				
<input type="checkbox"/> SEARCH FEE <small>(37 CFR 1.16(k), (l), or (m))</small>		N/A	N/A		N/A			N/A				
<input type="checkbox"/> EXAMINATION FEE <small>(37 CFR 1.16(o), (p), or (q))</small>		N/A	N/A		N/A			N/A				
TOTAL CLAIMS <small>(37 CFR 1.16(i))</small>		minus 20 =	*		X \$ =			X \$ =				
INDEPENDENT CLAIMS <small>(37 CFR 1.16(h))</small>		minus 3 =	*		X \$ =			X \$ =				
<input type="checkbox"/> APPLICATION SIZE FEE <small>(37 CFR 1.16(s))</small>		If the specification and drawings exceed 100 sheets of paper, the application size fee due is \$250 (\$125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).										
<input type="checkbox"/> MULTIPLE DEPENDENT CLAIM PRESENT <small>(37 CFR 1.16(j))</small>												
* If the difference in column 1 is less than zero, enter "0" in column 2.												
<b>APPLICATION AS AMENDED – PART II</b>												
(Column 1)			(Column 2)		SMALL ENTITY			OR			OTHER THAN SMALL ENTITY	
AMENDMENT	<b>02/24/2009</b>		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE (\$)	ADDITIONAL FEE (\$)		RATE (\$)	ADDITIONAL FEE (\$)	
	Total <small>(37 CFR 1.16(j))</small>		* 21	Minus	** 21	= 0	X \$ =		OR	X \$52=	0	
	Independent <small>(37 CFR 1.16(h))</small>		* 3	Minus	***3	= 0	X \$ =		OR	X \$220=	0	
	<input type="checkbox"/> Application Size Fee <small>(37 CFR 1.16(s))</small>											
	<input type="checkbox"/> FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM <small>(37 CFR 1.16(j))</small>											
TOTAL ADD'L FEE								OR	TOTAL ADD'L FEE			<b>0</b>
AMENDMENT			CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE (\$)	ADDITIONAL FEE (\$)		RATE (\$)	ADDITIONAL FEE (\$)	
	Total <small>(37 CFR 1.16(j))</small>		*	Minus	**	=	X \$ =		OR	X \$ =		
	Independent <small>(37 CFR 1.16(h))</small>		*	Minus	***	=	X \$ =		OR	X \$ =		
	<input type="checkbox"/> Application Size Fee <small>(37 CFR 1.16(s))</small>											
	<input type="checkbox"/> FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM <small>(37 CFR 1.16(j))</small>											
TOTAL ADD'L FEE								OR	TOTAL ADD'L FEE			
* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.												
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".												
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".												
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.												

Legal Instrument Examiner:  
/PATRICIA WARNER/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/354,569	02/14/2006	Ping Pan	HAMMP008	6912
21912	7590	11/20/2008	EXAMINER	
VAN PELT, YI & JAMES LLP 10050 N. FOOTHILL BLVD #200 CUPERTINO, CA 95014			LIU, SIMING	
			ART UNIT	PAPER NUMBER
			4145	
			MAIL DATE	DELIVERY MODE
			11/20/2008	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	11/354,569	PAN, PING	
	<b>Examiner</b>	<b>Art Unit</b>	
	SIMING LIU	4145	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1)  Responsive to communication(s) filed on 14 February 2006.
- 2a)  This action is **FINAL**.                      2b)  This action is non-final.
- 3)  Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4)  Claim(s) 1-21 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5)  Claim(s) \_\_\_\_\_ is/are allowed.
- 6)  Claim(s) 1-21 is/are rejected.
- 7)  Claim(s) \_\_\_\_\_ is/are objected to.
- 8)  Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9)  The specification is objected to by the Examiner.
- 10)  The drawing(s) filed on \_\_\_\_\_ is/are: a)  accepted or b)  objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11)  The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12)  Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
a)  All    b)  Some \*    c)  None of:
- 1.  Certified copies of the priority documents have been received.
  - 2.  Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - 3.  Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1)  Notice of References Cited (PTO-892)
- 2)  Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3)  Information Disclosure Statement(s) (PTO/SB/08)  
    Paper No(s)/Mail Date \_\_\_\_\_
- 4)  Interview Summary (PTO-413)  
    Paper No(s)/Mail Date \_\_\_\_\_
- 5)  Notice of Informal Patent Application
- 6)  Other: \_\_\_\_\_

## DETAILED ACTION

### *Specification*

1. The abstract of the disclosure is objected to because the title should not be included on the same sheet with the abstract. Correction is required. See MPEP § 608.01(b).
2. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

There is no 101 issue for claims 1-10. Because inherently there is a device which performs the steps recites in claims 1-10, the device is mentioned in the specification at line 8 of paragraph 0015, as a processor.

### ***Claim Rejections - 35 USC § 101***

3. Claims 17-21 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. "A computer program product" recites in claim 17-21 refer to computer readable medium in the specification. In the specification,

computer readable medium include “electronic communication link” which can be a signal. Signal is considered as non-statutory subject matter.

***Claim Rejections - 35 USC § 103***

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. **Claims 1-4, 7, 11-12, 15, 17, 20** are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang, US 2003/0117950 A1 in view of admission of prior art (AAPA), further in view of Blanchet, US 2004/0133692 A1.

1. Regarding **claim 1**, Huang teaches a method of providing protection to network traffic (Huang, page 2, para 0015, lines 1-6), comprising:  
sending (Huang, page 2, para 0016, right column, line 1: “receiving a request to set up”. There must be sending, thus receiving can happen. It’s inherent.) a ... protection configuration parameter (Huang, page 2, para 0016, right column, lines 2-4: “the request specifying a required protection bandwidth for the label switched path segment”, “required protection bandwidth” is equivalent to a protection configuration parameter) for configuring a standby ... between a source node and a destination node (Huang, page

2, para 0016, right column, lines 4-5: “and determining a backup route to the tail end node”);

...protection (Huang, page 2, para 0016, right column, lines 2-4: “the request specifying a required protection bandwidth for the label switched path segment”)...; and in the event that ... protection configuration parameters has been accepted by the destination node, using the standby ... (Huang, page 2, para 0016, right column, lines 6-14: In response of the request of setting up a label switched path segment over a direct connection between two nodes, a backup route is also being determined);

wherein the standby ... is configured based at least in part on the ... protection configuration parameter (Huang, page 2, para 0016, right column, lines 8-12: “The method also includes signaling to reserve the required protection bandwidth along the backup route, receiving confirmation of reservation of the required protection bandwidth and generating a backup connection map”, required protection bandwidth as a configuration parameter is a major factor in the backup path forming).

Huang doesn't expressly teach that Pseudowire.

AAPA teaches Pseudowire (AAPA, background of the inventions, para 0002, lines 3-4, lines 8-10: “A Pseudowire (PW) refers to an emulation of a native service over a network”).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to implement Pseudowire as a type of network service. The reason is that Pseudowire can emulate the operation of a “transparent wire” carrying the native

service. The method of modifying the system of Huang was within the ordinary ability of one of ordinary skill in the art based on the teachings of AAPA.

Huang in view of AAPA doesn't expressly teach that receiving a configuration acknowledgement indicating whether the configuration parameter has been accepted by the destination node.

Blanchet teaches that receiving a configuration acknowledgement indicating whether the configuration parameter has been accepted by the destination node (Blanchet, page 4, para 0035, lines 2-4).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the system to send an ACK indicating the acceptance of the configuration parameters. The rationale is that by sending out ACK indicating the acceptance of the configuration parameter makes the system more reliable. The method of modifying the system of Huang in view of AAPA was within the ordinary ability of one of ordinary skill in the art based on the teachings of Blanchet.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, AAPA and Blanchet to obtain the invention as specified in claim 1.

2. Regarding **claim 2**, Huang in view of AAPA and Blanchet as applied in claim 1 above teaches a method as recited in Claim 1, wherein the standby (Huang, page 2, para 0016, right column, lines 5: "backup" is equivalent to standby in the context) Pseudowire (AAPA, para 0002, lines 1-2) is configured to provide protection to at least



one primary (Huang, page 1, para 0008, lines 2-4) Pseudowire (previous discussed).

3. Regarding **claim 3**, Huang in view of AAPA and Blanchet as applied in claim 1 above teaches a method as recited in Claim 1, wherein the standby Pseudowire (previous discussed in claim 2) is configured to provide protection to at least one primary Pseudowire (previous discussed in claim 2), and in the event that the primary Pseudowire (previous discussed) fails to transfer network traffic (Huang, page 2, para 0010, line 5: "when a fault is discovered in a single link between two nodes along a path"), switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire (Huang, page 2, para 0010, lines 9-10: "switches the traffic that was using the connection to the alternate path", the limitation "Pseudowire" has been discussed).

4. Regarding **claim 4**, Huang in view of AAPA and Blanchet as applied in claim 1 above teaches a method as recited in Claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections (Huang, page 4, para 0040, lines 12-14: "The backup route may, for instance, be selected from a table of routes that have been pre-computed to connect the head end node 102A to the tail end node 102B"; page 4, para 0040, right column, lines 1-2: "a backup route can be determined instantaneously by the head end node 102A given information about the current state of the network 100").

5. Regarding **claim 7**, Huang in view of AAPA and Blanchet as applied in claim 1 above teaches a method as recited in Claim 1, wherein the Pseudowire protection configuration parameter (previous discussed) includes a protection scheme (Huang, page 2, para 0011, lines 9-10: "protection scheme").

6. Regarding **claim 10**, Huang in view of AAPA and Blanchet as applied in claim 1 above teaches a method as recited in Claim 1, wherein the Pseudowire protection configuration parameter (previous discussed) is established using the Label Distribution Protocol (LDP) (Huang, page 1, column 0005, right column, last 3 lines).

7. Regarding **claim 11**, a system for providing protection to network traffic, comprising:

a processor (Blanchet, Fig. 2, element 50: It's inherent, since all computers have at least one processor) configured to:

send a Pseudowire protection configuration parameter for configuring a

standby Pseudowire between a source node and a destination node; and

receive a Pseudowire configuration acknowledgement indicating whether

the Pseudowire protection configuration parameter has been accepted by the

destination node; and in the event that the Pseudowire protection configuration

parameter has been accepted by the destination node, use the standby Pseudowire;

wherein the standby Pseudowire is configured based at least in part on the

Pseudowire protection configuration parameter; and a memory coupled to the

processor, configured to provide the processor with instructions (Blanchet, Fig. 2, element 50: it's inherent since all computers have a memory coupled to a processor, and provide instructions with processor). (All of the remaining limitations have been discussed in claim 1)

8. Regarding **claim 12**, a system as recited in Claim 11, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire (All of the remaining limitations have been discussed in claim 2).

9. Regarding **claim 15**, a system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a protection scheme (All of the limitations have been discussed in claim 7).

10. Regarding **claim 17**, a computer program product (Huang, page 2, para 0013: "the 'gold' level of service protection" is a computer program product) for configuring a Pseudowire between a source node and a destination node, the computer program product (previous discussed) being embodied in a computer readable medium and comprising computer instructions (Blanchet, Fig. 2, element 50: it's inherent, since all computers have memory and can give out computer instructions) for:  
sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node;  
receiving a Pseudowire configuration acknowledgement indicating whether the

Pseudowire protection configuration parameter has been accepted by the destination node; and in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire; wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter (All of the remaining limitations have been discussed in claim 1).

11. Regarding **claim 20**, Huang in view of AAPA, Blachet as applied in claim above teaches a computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a protection scheme (All of the remaining limitations have been discussed in claim 7).

12. **Claims 5, 13, 18** are rejected under 35 U.S.C. 103(a) as being unpatentable over under 35 U.S.C. 103(a) as being unpatentable over Huang, US 2003/0117950 A1 in view of admission of prior art (AAPA), further in view of Blanchet, US 2004/0133692 A1, further in view of Cruz, US 2006/0046658 A1.

13. Regarding **claim 5**, Huang in view of AAPA and Blanchet as applied in claim above teaches a method as recited in Claim 1, wherein the Pseudowire (AAPA, para 0002, lines 1-2) protection configuration parameter (Huang, page 2, para 0016, right column, lines 2-4: "the request specifying a required protection bandwidth for the label switched path segment", "required protection bandwidth" is equivalent to a protection configuration parameter) includes ...

Huang in view of AAPA and Blanchet doesn't expressly teach that a domain type.

Cruz teaches a domain type (Cruz, page 1, para 0017, line 2: According to the specification of the application, domain type is about whether the network is either multi-hop or single hop).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include domain type. The reason is that by including domain type in the configuration parameter, it would be more accurate to select a desired standby path, given that you have more information about the network. The method of change the configuration parameter by including the domain type of Huang in view of AAPA and Blanchet was within the ordinary ability of one of ordinary skill in the art based on the teachings of Cruz.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, AAPA, Blanchet and Cruz to obtain the invention as specified in claim 5.

14. Regarding **claim 13**, a system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a domain type (All of the remaining limitations have been discussed in claim 5).

15. Regarding **claim 18**, a computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a domain type (All of the remaining limitations have been discussed in claim 5).

16. **Claims 6, 14, 19** are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang, US 2003/0117950 A1 in view of admission of prior art (AAPA), further in view of Blanchet, US 2004/0133692 A1, further in view of Rathunde, US 6,574,477 B1.

Regarding **claim 6**, a method as recited in Claim 1, wherein the Pseudowire protection configuration parameter (previous discussed) includes ...

Huang in view of AAPA and Blanchet doesn't expressly teach that a protection type.

Rathunde teaches a protection type (Rathunde, col 9, line 3: "type of standby mode", according to the specification of the application, protection type just means what type of standby mode).

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include a protection type. The reason is that by including protection type in the configuration parameter, it would be more accurate to select a desire standby path, given that you have more information about the network. The method of change the configuration parameter by including the protection type of Huang in view of AAPA and Blanchet was within the ordinary ability of one of ordinary skill in the art based on the teachings of Rathunde.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, AAPA, Blanchet and Rathunde to obtain the invention as specified in claim 6.

17. Regarding **claim 14**, a system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a protection type (All of the limitations have been discussed in claim 6).

18. Regarding **claim 19**, a computer program product (previous discussed) as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a protection type (All of the remaining limitations have been discussed in claim 6).

19. **Claims 8-9, 16, 21** are rejected under 35 U.S.C. 103(a) as being unpatentable over Huang, US 2003/0117950 A1 in view of admission of prior art (AAPA), further in view of Blanchet, US 2004/0133692 A1, further in view of Saleh, US 7,200,104 B2.

20. Regarding **claim 8**, a method as recited in Claim 1, wherein the Pseudowire protection configuration parameter (previous discussed) includes a ...

Huang in view of AAPA and Blanchet doesn't expressly teach that a priority.

Saleh teaches a priority (Saleh, col 3, line 38: "restoration priority level").

At the time of the invention, it would have been obvious to a person of ordinary skill in the art to modify the configuration parameter to include priority. The reason is that by including domain type in the configuration parameter, it would be more accurate

to select a desire standby path, given that you have more information about the network. The method of change the configuration parameter by including the domain type of Huang in view of AAPA and Blanchet was within the ordinary ability of one of ordinary skill in the art based on the teachings of Saleh.

Therefore, it would have been obvious to one of the ordinary skill in the art to combine the teachings of Huang, AAPA, Blanchet and Saleh to obtain the invention as specified in claim 8.

21. Regarding **claim 9**, a method as recited in Claim 1, further including determining whether to preempt existing traffic on the standby Pseudowire, the determination being based at least in part on a priority associated with the standby Pseudowire (Saleh, col 3, lines 1-5).

22. Regarding **claim 16**, a system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a priority (All of the limitations have been discussed in claim 8).

23. Regarding **claim 21**, a computer program product (previous discussed) as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a priority (All of the remaining limitations have been discussed in claim 8).



**Conclusion**

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SIMING LIU whose telephone number is (571)270-3859. The examiner can normally be reached on Monday-Friday 8:30am-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Pankaj Kumar can be reached on 571-272-3011. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

SL  
/Pankaj Kumar/  
Supervisory Patent Examiner, Art Unit 4145

<b>Notice of References Cited</b>	Application/Control No. 11/354,569	Applicant(s)/Patent Under Reexamination PAN, PING	
	Examiner SIMING LIU	Art Unit 4145	Page 1 of 1

**U.S. PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A US-2003/0117950	06-2003	Huang, Gail G.	370/220
*	B US-7,200,104	04-2007	Saleh et al.	370/216
*	C US-6,574,477	06-2003	Rathunde, Dale Frank	455/453
*	D US-2004/0133692	07-2004	Blanchet et al.	709/230
*	E US-2006/0046658	03-2006	Cruz et al.	455/067.11
	F US-			
	G US-			
	H US-			
	I US-			
	J US-			
	K US-			
	L US-			
	M US-			

**FOREIGN PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N				
	O				
	P				
	Q				
	R				
	S				
	T				

**NON-PATENT DOCUMENTS**

*	Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)				
	U				
	V				
	W				
	X				


\*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)  
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

<b>Index of Claims</b>  	<b>Application/Control No.</b> 11354569	<b>Applicant(s)/Patent Under Reexamination</b> PAN, PING
	<b>Examiner</b> SIMING LIU	<b>Art Unit</b> 4145

✓	<b>Rejected</b>	-	<b>Cancelled</b>	N	<b>Non-Elected</b>	A	<b>Appeal</b>
=	<b>Allowed</b>	÷	<b>Restricted</b>	I	<b>Interference</b>	O	<b>Objected</b>

Claims renumbered in the same order as presented by applicant
  CPA
  T.D.
  R.1.47

CLAIM		DATE								
Final	Original	10/30/2008								
	1	✓								
	2	✓								
	3	✓								
	4	✓								
	5	✓								
	6	✓								
	7	✓								
	8	✓								
	9	✓								
	10	✓								
	11	✓								
	12	✓								
	13	✓								
	14	✓								
	15	✓								
	16	✓								
	17	✓								
	18	✓								
	19	✓								
	20	✓								
	21	✓								

<b>Search Notes</b>  	<b>Application/Control No.</b>  11354569	<b>Applicant(s)/Patent Under Reexamination</b>  PAN, PING
	<b>Examiner</b>  SIMING LIU	<b>Art Unit</b>  4145

SEARCHED			
Class	Subclass	Date	Examiner
370	216, 225, 228	10/30/2008	/SL/
709	220	10/30/2008	/SL/

SEARCH NOTES		
Search Notes	Date	Examiner
East Class search	10/30/2008	/SL/
Palm inventor name search	10/30/2008	/SL/
Consulted 101 issues with Peng, John	11/10/2008	/SL/

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner

--	--

## EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S5	115	pseudowire	US-PGPUB; USPAT	OR	ON	2008/10/08 12:53
S6	0	pseudowire and tele	US-PGPUB; USPAT	OR	ON	2008/10/08 12:53
S7	217	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2008/10/08 13:07
S8	9	S7 with protection	US-PGPUB; USPAT	OR	ON	2008/10/08 13:08
S9	4	S7 with protection and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 13:09
S10	1	"20040223498".pn.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:07
S11	0	(pseudowire or pseudo-wire) and initiliz\$5	US-PGPUB; USPAT	OR	ON	2008/10/08 14:14
S12	133	(pseudowire or pseudo-wire) and initi\$5	US-PGPUB; USPAT	OR	ON	2008/10/08 14:15
S13	51	(pseudowire or pseudo-wire) and initi\$5 and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:15
S14	2193	(370/216,225,228).ccls.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:17
S15	6	(370/216,225,228).ccls. and S7	US-PGPUB; USPAT	OR	ON	2008/10/08 14:23
S16	3	(709/220).ccls. and S7	US-PGPUB; USPAT	OR	ON	2008/10/08 14:26
S17	31	((PING) near2 (PAN)).INV.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:32
S18	2	((PING) near2 (PAN)).INV. and pseudowire	US-PGPUB; USPAT	OR	ON	2008/10/08 14:33
S19	2	((PING) near2 (PAN)).INV. and (pseudowire).clm.	US-PGPUB; USPAT	OR	ON	2008/10/08 14:33
S20	66	S7 and (primary)	US-PGPUB; USPAT	OR	ON	2008/10/08 14:38
S21	23	S7 and (primary) and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:39
S22	75	S7 and (config\$7) and @ad< "20050214"	US-PGPUB; USPAT	OR	ON	2008/10/08 14:44
S23	2	TDM pseudowire	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:09
S24	106	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol))	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:16

S26	43	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:17
S27	11	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (standby or backup)	US-PGPUB; USPAT	ADJ	ON	2008/10/08 15:19
S28	9	(pseudowire or (pseudo wire) or pseudo-wire) and (LDP or (Label Distribution protocol)) and @ad<"20050214" and (primary or main) and (secondly or backup or standby)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 09:00
S29	43	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:34
S30	14	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:35
S31	0	(pseudowire or (pseudo wire) or pseudo-wire) same (config\$7 with parameter) same (destination near5 node)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:38
S32	18	(pseudowire or (pseudo wire) or pseudo-wire) and ((config\$7) same (destination near5 node))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:38
S33	1	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 with acknowledgement)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:41
S34	0	(pseudowire or (pseudo wire) or pseudo-wire) and (config same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:44
S35	8	(pseudowire or (pseudo wire) or pseudo-wire) and (config\$7 same (acknowledgement or ack))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:44
S36	370	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge))	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:47
S37	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S33) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:48

S38	218	pseudowire or pseudo-wire	US-PGPUB; USPAT	OR	ON	2008/10/09 10:49
S39	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S38) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:49
S40	275	pseudowire or pseudo-wire or (pseudo wire)	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:49
S41	0	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and (S40) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:50
S42	233	(send\$5 with config\$7 with parameter) and (receiv\$5 with (ack or acknowledge)) and @ad<"20050214"	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:50
S43	27	S40 and initialization	US-PGPUB; USPAT	ADJ	ON	2008/10/09 10:54
S44	3434011	(link or route or path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2008/10/29 09:26
S46	1334019	(fail\$5 or (stop\$1 working))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:27
S47	3640734	(alter\$7 or backup or standby)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:28
S48	29788561	@ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:28
S49	6386553	(pick\$5 or select\$5 or choos\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:30
S50	409	(S44 near7 S46) with (S49 near7 S47 near7 S44) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:31
S51	238	(S44 near7 S46) with (S49 near7 S47 near7 S44) and S48 and (priority or bandwidth)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:38
S52	25	(S44 near7 S46) with (S49 near7 S47 near7 S44) same (priority or bandwidth or parameter) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:43

S53	2289	S47 with config\$7 with (primary near7 S47)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:49
S54	159	(S47 near5 S44) with config \$7 with (primary near7 S47)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:54
S55	175	(S47 near5 S44) with config \$7 with (primary near7 S44)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:54
S56	111	(S47 near5 S44) with config \$7 with (primary near7 S44) and S48	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 09:56
S57	7	09/859166	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 10:26
S58	33	(restoration scheme) and ("1:N")	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 10:50
S59	4	(restoration scheme) and (priority) and (standby mode)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 12:42
S60	18	(restoration scheme) and (priority) and (config\$7 near5 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:27
S61	3706723	(send\$7 or transmit\$5)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:30
S62	0	(source node) with S61 with (config\$7 near3 parameter \$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:31
S63	5	(source) with S61 with (config\$7 near3 parameter \$1) with (destin\$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:32
S64	569	(source) with S61 with (parameter\$1) with (destin \$7)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:33
S65	54	(source node) with S61 with (parameter\$1) with (destin \$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:33



S66	2959	(ack or acknowledgement) and (config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:53
S67	0	(ack or acknowledgement) same (config\$7 parameter \$1) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:53
S68	20807	(config\$7 parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:54
S69	52906	(ack or acknowledgement) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:54
S70	29	(ack or acknowledgement) and (restoration scheme) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 13:55
S71	137	S61 with (parameter\$1) with (destin\$7 node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:19
S72	0	handshaking with (restoration scheme)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:29
S73	10549	handshaking and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:29
S74	759	handshaking and @ad<"20050214" and (S44 with S46)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:30
S75	2	"6553034".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:32
S76	108	(virtual path) and ((protection or restoration) near5 scheme) and priority	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:34
S77	103	(virtual path) and ((protection or restoration) near5 scheme) and priority and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:34
S78	3479	(protection or restoration) near5 parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:51

S79	2628	((protection or restoration) near5 parameter and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:52
S80	6	((protection or restoration) near5 parameter) with (S61) and (destin\$7 near3 node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:53
S81	26	((protection or restoration) near5 parameter) and (handshaking) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 16:55
S82	73	((protection or restoration) near5 parameter) and (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:04
S83	0	((protection or restoration) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:04
S84	0	((protection or restoration or config\$7) near5 parameter) wotj (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S85	15	((protection or restoration or config\$7) near5 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S86	4	((protection or restoration or config\$7) near2 parameter) with (destination node) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:05
S87	651	handshaking and @ad<"20050214" and (config\$7 parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/29 17:10
S88	0	receiving acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:27
S89	0	receiv\$7 acknowledgement indicat\$7 parameter accept \$7 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:27
S90	276	receiv\$7 acknowledgement destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:28
S91	0	receiv\$7 acknowledgement (parameter accept\$5 destination node)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:40

S92	5	receiv\$7 acknowledgement accept\$3 destination node	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:45
S93	7	receiv\$7 acknowledgement parameter accept\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 09:48
S94	2	"20030117950".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:20
S95	771	(domain type) with (parameter)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:25
S96	3	(parameter) near5 includ\$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:26
S97	3	(parameter\$1) near5 includ \$5 near5 (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:27
S98	64	(parameter\$1) with (domain adj type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	WITH	ON	2008/10/30 12:27
S99	0	(domain type) with (single- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:28
S100	0	(domain type) with (single near5 hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:29
S101	64	(domain type) with parameter	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:29
S102	179	(single-hop) same (multi- hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:32
S103	9	(single-hop) same (multi- hop) same (parameter\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:33
S104	0	field with indica\$7 with ((single-hop) same (multi- hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:34

S105	147	field with indica\$7 with (topology)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:34
S106	6	field with indica\$7 with (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:36
S107	179	(single-hop) same (multi-hop)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:37
S108	10	(field or parameter) same ((single-hop) same (multi-hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:37
S109	283	((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:40
S111	21	(parameter or field) same S109	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:41
S112	0	S109 same (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:44
S113	134	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:46
S114	0	parameter with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:48
S115	0	(field or parameter) with indicat\$5 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:49
S116	0	(field or parameter) with (show\$3 or indicat\$5) same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:49
S117	0	(field or parameter) with (domain type) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50

S118	68	(field or parameter) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S119	14	(field or parameter) with (indicat\$5 or show\$5) with (domain type) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 12:50
S120	1	(protection type) and (standby path)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:44
S121	2636	(hot or warm or cold) near3 standby	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:46
S122	283	(hot and cold) same standby and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:47
S123	51	(hot and cold) and (parameter with standby) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:48
S124	20	(field with indicat\$5 with (standby mode)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:49
S126	696	config\$9 with (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:50
S127	406	config\$9 near7 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:51
S128	324	config\$9 near5 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:51
S129	194	config\$9 near3 (standby mode) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:52
S130	7	config\$9 near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 13:54
S131	9	type with (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:01

S132	4	type near3 (standby mode) and @ad<"20050214" and (hot and cold)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:01
S133	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (domain type)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:38
S135	134	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:39
S136	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:43
S137	126	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:43
S138	11	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) same (parameter or field) and @ad<"20050214"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:44
S139	0	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 netowrk)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:53
S140	136	((single hop) or (single-hop)) same ((multi-hop) or (multi hop)) and (type near5 network)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:53
S141	0	(paramete or field or bit) with indicat\$7 with ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:55
S142	2	(paramete or field or bit) with indicat\$7 same ((single hop) or (single-hop)) same ((multi-hop) or (multi hop))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	ADJ	ON	2008/10/30 14:55

10/30/2008 6:23:26 PM

C:\Documents and Settings\slu3\My Documents\EAST\Workspaces\11354569.wsp



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
 United States Patent and Trademark Office  
 Address: COMMISSIONER FOR PATENTS  
 P.O. Box 1450  
 Alexandria, Virginia 22313-1450  
 www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 6912

SERIAL NUMBER	FILING or 371(c) DATE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.		
11/354,569	02/14/2006	362	4145	HAMMP008		
<b>RULE</b>						
<b>APPLICANTS</b> Ping Pan, San Jose, CA; <b>** CONTINUING DATA *****</b> This appln claims benefit of 60/653,065 02/14/2005 <b>** FOREIGN APPLICATIONS *****</b> <b>** IF REQUIRED, FOREIGN FILING LICENSE GRANTED **</b> 03/17/2006						
Foreign Priority claimed <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No 35 USC 119(a-d) conditions met <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No Verified and Acknowledged <u>/SIMING LIU/</u> Examiner's Signature		<input type="checkbox"/> Met after Allowance Initials	<b>STATE OR COUNTRY</b> CA	<b>SHEETS DRAWINGS</b> 7	<b>TOTAL CLAIMS</b> 21	<b>INDEPENDENT CLAIMS</b> 3
<b>ADDRESS</b> VAN PELT, YI & JAMES LLP 10050 N. FOOTHILL BLVD #200 CUPERTINO, CA 95014 UNITED STATES						
<b>TITLE</b> Pseudowire protection						
<b>FILING FEE RECEIVED</b> 1180	FEES: Authority has been given in Paper No. _____ to charge/credit DEPOSIT ACCOUNT No. _____ for following:		<input type="checkbox"/> All Fees <input type="checkbox"/> 1.16 Fees (Filing) <input type="checkbox"/> 1.17 Fees (Processing Ext. of time) <input type="checkbox"/> 1.18 Fees (Issue) <input type="checkbox"/> Other _____ <input type="checkbox"/> Credit			

Day : Wednesday

Date: 10/8/2008  
Time: 14:28:37



Inventor Name Search Result Office of Public Affairs

Your Search was:

Last Name = PAN  
First Name = PING

Application#	Patent#	PG Pub#	Status	Date Filed	Title	Examiner Name
<a href="#">60653065</a>	Not Issued	20070163769	159	02/14/2005	PSEUDO WIRE PROTECTION	
<a href="#">11787664</a>	Not Issued	20070163769	041	04/16/2007	HYBRID DATA SWITCHING FOR EFFICIENT PACKET PROCESSING	DUONG,FRANK
<a href="#">61064357</a>	Not Issued	20070163769	020	02/29/2008	METHOD AND APPARATUS FOR EVENT-PROFILE-BASED INTERACTIVE VIDEO	
<a href="#">12101245</a>	Not Issued	20070163769	025	04/11/2008	GRACEFUL RESTART FOR USE IN NODES EMPLOYING LABEL SWITCHED PATH SIGNALING PROTOCOLS	
<a href="#">60360786</a>	Not Issued	20070163769	159	02/28/2002	DETECTING DATA PLANE LIVELINESS IN MPLS	
<a href="#">10095000</a>	<a href="#">7359377</a>	20070163769	150	03/11/2002	GRACEFUL RESTART FOR USE IN NODES EMPLOYING LABEL SWITCHED PATH SIGNALING PROTOCOLS	HALIYUR,VENKATI



<u>10142730</u>	Not Issued	20070163769	071	05/08/2002	AGGREGATING END-TO-END QOS SIGNED PACKET FLOWS THROUGH LABEL SWITCHED PATHS	SAM,PHIRIN
<u>10179927</u>	<u>7336615</u>	20070163769	150	06/25/2002	DETECTING DATA PLANE LIVELINES IN CONNECTIONS SUCH AS LABEL-SWITCHED PATHS	HAILE,FEEN
<u>60444440</u>	Not Issued	20070163769	159	02/03/2003	PSEUDO-WIRE ADMISSION CONTROL EXTENSION	
<u>60698893</u>	Not Issued	20070163769	159	07/12/2005	SUPPORTING PSEUDO-WIRES IN SUB-IP ACCESS NETWORKS	
<u>11184171</u>	Not Issued	20070163769	041	07/19/2005	METHOD AND APPARATUS FOR INTERFACING APPLICATIONS TO LCAS FOR EFFICIENT SONET TRAFFIC FLOW CONTROL	KAMARA,MOHAME
<u>60726115</u>	Not Issued	20070163769	159	10/12/2005	IMS-BASED NETWORK CONVERGENCE WITH THE HSX	
<u>60725038</u>	Not Issued	20070163769	159	10/07/2005	APPLICATION WIRE: MAPPING APPLICATION STREAMS TO PSEUDO-WIRES	
<u>11890308</u>	Not Issued	20070163769	030	08/03/2007	GLOBAL IP-BASED SERVICE-ORIENTED NETWORK ARCHITECTURE	PATEL,JAYANTI
<u>60996580</u>	Not Issued	20070163769	020	11/26/2007	METHOD AND APPARATUS FOR	

					INTERACTIVE VIDEO ADVERTISEMENT OVER THE INTERNET	
<u>60990197</u>	Not Issued	20070163769	160	01/01/0001	METHOD AND APPARATUS FOR INTERACTIVE VIDEO ADVERTISEMENT OVER THE INTERNET	
<u>10757528</u>	<u>6985488</u>	20070163769	150	01/15/2004	METHOD AND APPARATUS FOR TRANSPORTING PACKET DATA OVER AN OPTICAL NETWORK	LEVITAN,DMITRY
<u>11354569</u>	Not Issued	20070163769	030	02/14/2006	PSEUDOWIRE PROTECTION	LIU,SIMING
<u>10769891</u>	<u>7417950</u>	20070163769	150	02/03/2004	METHOD AND APPARATUS FOR PERFORMING DATA FLOW INGRESS/EGRESS ADMISSION CONTROL IN A PROVIDER NETWORK	MURPHY,RHONDA
<u>60440313</u>	Not Issued	20070163769	159	01/15/2003	METHOD AND APPARATUS FOR TRANSPORTING LAYER-2 TRAFFIC OVER SONET/SDH NETWORKS	
<u>11580530</u>	Not Issued	20070163769	030	10/12/2006	CONTROL PLANE TO DATA PLANE BINDING	MIAN,OMER
<u>61083829</u>	Not Issued	20070163769	020	07/25/2008	AUTO PROVISIONING FOR METRO ETHERNET NETWORK ELEMENTS	
<u>11543727</u>	Not	20070163769	030	10/05/2006	APPLICATION	HASPEL,AMY

	Issued				WIRE	
<u>60835794</u>	Not Issued	20070163769	159	08/04/2006	GLOBAL IP-BASED SERVICE-ORIENTED NETWORK ARCHITECTURE OVERVIEW AND IMS USER CASE	
<u>11486389</u>	Not Issued	20070163769	030	07/12/2006	LIGHTWEIGHT CONTROL-PLANE SIGNALING FOR AGGREGATION DEVICES IN A NETWORK	PATEL,JAYANTI
<u>11486432</u>	Not Issued	20070163769	030	07/12/2006	PROXIES FOR PSEUDO-WIRE ALLOCATION AND DISTRIBUTION	PATEL,JAYANTI
<u>60301050</u>	Not Issued	20070163769	159	06/25/2001	DETECTING DATA PLANE LIVELINESS IN RSVP-TE	
<u>60299813</u>	Not Issued	20070163769	159	06/19/2001	GRACEFUL RESTART MECHANISM FOR RSVP-TE	
<u>60792078</u>	Not Issued	20070163769	159	04/14/2006	HYBRID SWITCHING METHOD FOR EFFICIENT PACKET PROCESSING	
<u>60589004</u>	Not Issued	20070163769	159	07/20/2004	METHOD AND APPARATUS FOR INTERFACING APPLICATIONS TO LCAS FOR EFFICIENT SONET TRAFFIC FLOW CONTROL	
<u>10869501</u>	Not Issued	20070163769	041	06/16/2004	PROTECTING CONNECTION TRAFFIC USING FILTERS	COULTER,KENNETH
<u>10357262</u>	Not Issued	20070163769	041	02/03/2003	DETECTING A LABEL-	ENGLAND,DAVID

					SWITCHED PATH OUTAGE USING ADJACENCY INFORMATION	
<a href="#">10365598</a>	Not Issued	20070163769	093	02/12/2003	DETECTING DATA PLANE LIVELINESS OF A LABEL-SWITCHED PATH	TSEGAYE,SABA
<a href="#">60444456</a>	Not Issued	20070163769	159	02/03/2003	DRY-MARTINI APPLICATIONS ON PALM	
<a href="#">10165643</a>	Not Issued	20070163769	161	06/07/2002	COMBINATION THERAPY FOR THE PREVENTION OR TREATMENT OF CANCER, INFLAMMATORY DISORDERS OR INFECTIOUS DISEASES IN A SUBJECT	LI,QIAN
<a href="#">60816863</a>	Not Issued	20070163769	159	06/28/2006	APPARATUS AND FILE FORMAT FOR TEXT WITH SYNCHRONIZED AUDIO	
<a href="#">11812133</a>	Not Issued	20070163769	030	06/15/2007	APPARATUS, METHOD, AND FILE FORMAT FOR TEXT WITH SYNCHRONIZED AUDIO	HERNDON,HEATHE

Inventor Search Completed: No Records to Display.

Search Another: Inventor

Last Name	First Name
<input type="text" value="PAN"/>	<input type="text" value="PING"/>

Enter both names for a faster result, even if it is only a few letters.

(To go back use Back button on your browser toolbar)



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor:	Ping Pan	Examiner:	Not Assigned
Application No.:	11/354,569	Art Unit:	2875
Filed:	February 14, 2006	Docket No.:	HAMMP008
Title:	PSEUDOWIRE PROTECTION		

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in a prepaid envelope addressed to: Mail Stop Missing Parts, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on:

5/11, 2006.

Elaine Nguyen  
Elaine Nguyen

RESPONSE TO NOTICE TO FILE MISSING PARTS

MAIL STOP MISSING PARTS

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

This is a Response to Notice to File Missing Parts mailed March 20, 2006.

Application Elements:

- Declaration
  - Newly executed (original or copy)
    - Copy from a prior application (37 CFR 1.63(d) for a continuation or divisional).  
The entire disclosure of the prior application from which a copy of the declaration is herein supplied is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.
    - Deletion of inventors Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).

Accompanying Application Parts:

- Information Disclosure Statement with Form PTO/SB/08
- Copies of IDS Citations
- Preliminary Amendment (New claims numbered after highest original claim in prior application.)
- Return Receipt Postcard
- Copy of Notice to File Missing Parts of NonProvisional Application
- Other:

Amendments

Cancel in this application original claims \_\_\_\_\_ of the prior application before calculating the filing fee. (At least one original independent claim must be retained.)

Fee Calculation (37 CFR § 1.16)

				Small Entity		Large Entity	
				Rate	Fee	Rate	Fee
Filing Fees				x \$150 = \$		OR	x \$300 = \$ 300
Search Fees				x \$250 = \$		OR	x \$500 = \$ 500
Examination Fees				x \$100 = \$		OR	x \$200 = \$ 200
CLAIMS	Filed.		Extra				
Total	21	Less 20	1	x \$25 = \$		OR	x \$50 = \$ 50.00
Independent	3	Less 3	-0-	x \$100 = \$		OR	x \$200 = \$
Multiple Dependent Claims					-0-	OR	x \$360 = \$
Declaration and Surcharge Fee				X \$65 = \$		OR	X \$130 = \$ 130
				TOTAL FILING FEE \$			TOTAL FILING FEE \$ <b>1180.00</b>

Applicants petition for an extension of time to respond under 37 CFR § 1.136(a) as follows:

	SMALL ENTITY		LARGE ENTITY	
	Rate	Add'l Fee	Rate	Add'l Fee
<input type="checkbox"/> Extension for Response within FIRST month	x \$60 = \$		OR	x \$120 = \$
<input type="checkbox"/> Extension for Response within SECOND month	x \$225 = \$		OR	x \$450 = \$
<input type="checkbox"/> Extension for Response within THIRD month	x \$510 = \$		OR	x \$1020 = \$
<input type="checkbox"/> Extension for Response within FOURTH month	x \$795 = \$		OR	x \$1590 = \$
<input type="checkbox"/> Extension for Response within FIFTH month	x \$1080 = \$		OR	x \$2160 = \$

Check No. 2332 for \$1180.00 is enclosed.

General Authorizations

Applicants hereby make and generally authorize any Petitions for Extensions of Time as needed for this or any subsequent filings. The Commissioner is also authorized to charge any extension fees under 37 CFR §1.17 as needed to Deposit Account No. 50-0685 (Order No. HAMMP008).

The Commissioner is given general authorization to charge any fees or to credit any overpayment during the pendency of this application to Deposit Account No. 50-0685 (Order No. HAMMP008).

Please send correspondence to the following address:

**Customer No. 21912**  
 VAN PELT, YI & JAMES LLP  
 10050 N. Foothill Blvd.  
 Suite 200  
 Cupertino, CA 95014

Date: 4-28-06

Clover Huang  
 Clover Huang  
 Reg. No. 55,285

JRW ✓



## UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
 United States Patent and Trademark Office  
 Address: COMMISSIONER FOR PATENTS  
 P.O. Box 1450  
 Alexandria, Virginia 22313-1450  
 www.uspto.gov

APPLICATION NUMBER	FILING OR 371 (c) DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NUMBER
11/354,569	02/14/2006	Ping Pan	HAMMP008

21912  
 VAN PELT, YI & JAMES LLP  
 10050 N. FOOTHILL BLVD #200  
 CUPERTINO, CA 95014

05/16/2006 SSITHIB1 00000024 11354569

01 FC:1011	300.00 OP
02 FC:1111	500.00 OP
03 FC:1311	200.00 OP
04 FC:1051	130.00 OP



CONFIRMATION NO. 6912  
 FORMALITIES  
 LETTER

Date Mailed: 03/20/2006

## NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

05 FC:1202 50.00 OP

FILED UNDER 37 CFR 1.53(b)

*Filing Date Granted*

**Items Required To Avoid Abandonment:**

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing. Applicant is given **TWO MONTHS** from the date of this Notice within which to file all required items and pay any fees required below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- The statutory basic filing fee is missing.  
*Applicant must submit \$ 300 to complete the basic filing fee for a non-small entity. If appropriate, applicant may make a written assertion of entitlement to small entity status and pay the small entity filing fee (37 CFR 1.27).*
- The oath or declaration is missing. *A properly signed oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.*  
*Note: If a petition under 37 CFR 1.47 is being filed, an oath or declaration in compliance with 37 CFR 1.63 signed by all available joint inventors, or if no inventor is available by a party with sufficient proprietary interest, is required.*

The applicant needs to satisfy supplemental fees problems indicated below.

The required item(s) identified below must be timely submitted to avoid abandonment:

- Additional claim fees of \$50 as a non-small entity, including any required multiple dependent claim fee, are required. Applicant must submit the additional claim fees or cancel the additional claims for which fees are due.
- To avoid abandonment, a surcharge (for late submission of filing fee, search fee, examination fee or oath or declaration) as set forth in 37 CFR 1.16(f) of \$130 for a non-small entity, must be submitted with the missing items identified in this letter.

**SUMMARY OF FEES DUE:**

Total additional fee(s) required for this application is \$1180 for a Large Entity

JUNIPER Exhibit 1003

App. 3, pg. 301

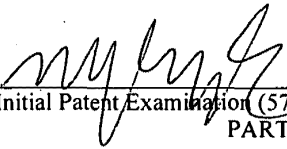
'652 File History 301

- \$300 Statutory basic filing fee.
- \$130 Surcharge.
  
- The application search fee has not been paid. Applicant must submit \$500 to complete the search fee.
- The application examination fee has not been paid. Applicant must submit \$200 to complete the examination fee for a large entity
  
- Total additional claim fee(s) for this application is \$50
  - \$50 for 1 total claims over 20.

Replies should be mailed to: Mail Stop Missing Parts  
Commissioner for Patents  
P.O. Box 1450  
Alexandria VA 22313-1450

---

*A copy of this notice **MUST** be returned with the reply.*

  
Office of Initial Patent Examination (571) 272-4000, or 1-800-PTO-9199, or 1-800-972-6382

PART 2 - COPY TO BE RETURNED WITH RESPONSE





**DECLARATION AND POWER OF ATTORNEY  
FOR ORIGINAL U.S. PATENT APPLICATION**

Attorney's Docket No. HAMMP008

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled: **PSEUDOWIRE PROTECTION**, the specification of which,

- (check one) 1.  is attached hereto.
- 2.  was filed on 2/14/2006 as  
U.S. Application No. 11/354,569  
and was amended on \_\_\_\_\_.
- 3.  was filed on \_\_\_\_\_ as  
International PCT Application No. \_\_\_\_\_  
and was amended on \_\_\_\_\_.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

*For Assigned Inventions:* I understand that the purpose of making this appointment is to permit prosecution of patent applications for the above-identified invention for the benefit of my assignee, and that this appointment does not create an attorney-client relationship between me and these appointees.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, CFR § 1.56.

I hereby claim foreign priority benefits under Title 35, United States code, § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have identified below, by checking the box, any foreign application for patent or inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed:

<b>Prior Foreign Application(s)</b>	<b>Priority Benefits Claimed?</b>
_____	<input type="checkbox"/> Yes <input type="checkbox"/> No
(Appl. No.)                      (Country)                      (Filing Date)	

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below:

<b>Prior Provisional Application(s)</b>
<u>60/653,065</u> <u>2/14/2005</u>
(Application No.)                      (Filing Date)

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s), or § 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

**Prior U.S. Application(s)**

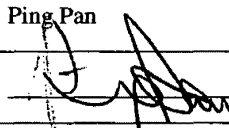
(Application No.)	(Filing Date)	(Status - patented, pending, abandoned)
-------------------	---------------	---

And I hereby appoint the attorneys and/or agents associated with Customer No. **21912** as my principal attorneys to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

**Please Direct all Correspondence To: Customer No. 21912**  
 VAN PELT, YI & JAMES LLP  
 10050 N. Foothill Blvd., Suite 200  
 Cupertino, CA 95014

**Direct Telephone Calls To: Clover Huang at telephone number (408) 973-2585**

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Full name of sole or first inventor	Ping Pan		
Signature of sole or first inventor		Date:	4/25, 2006
Residence: City	San Jose	State:	CA
		Citizenship:	United States of America
Mailing Address	640 Clyde Court, Mountain View, CA 94043		



## UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
 United States Patent and Trademark Office  
 Address: COMMISSIONER FOR PATENTS  
 P.O. Box 1450  
 Alexandria, Virginia 22313-1450  
 www.uspto.gov

APPLICATION NUMBER	FILING OR 371 (c) DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NUMBER
11/354,569	02/14/2006	Ping Pan	HAMMP008

21912  
 VAN PELT, YI & JAMES LLP  
 10050 N. FOOTHILL BLVD #200  
 CUPERTINO, CA 95014

**CONFIRMATION NO. 6912**  
**FORMALITIES**  
**LETTER**

Date Mailed: 03/20/2006

## NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

FILED UNDER 37 CFR 1.53(b)

*Filing Date Granted*

### Items Required To Avoid Abandonment:

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing. Applicant is given **TWO MONTHS** from the date of this Notice within which to file all required items and pay any fees required below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- The statutory basic filing fee is missing.  
*Applicant must submit \$ 300 to complete the basic filing fee for a non-small entity. If appropriate, applicant may make a written assertion of entitlement to small entity status and pay the small entity filing fee (37 CFR 1.27).*
- The oath or declaration is missing. *A properly signed oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.*  
*Note: If a petition under 37 CFR 1.47 is being filed, an oath or declaration in compliance with 37 CFR 1.63 signed by all available joint inventors, or if no inventor is available by a party with sufficient proprietary interest, is required.*

The applicant needs to satisfy supplemental fees problems indicated below.

The required item(s) identified below must be timely submitted to avoid abandonment:

- Additional claim fees of \$50 as a non-small entity, including any required multiple dependent claim fee, are required. Applicant must submit the additional claim fees or cancel the additional claims for which fees are due.
- To avoid abandonment, a surcharge (for late submission of filing fee, search fee, examination fee or oath or declaration) as set forth in 37 CFR 1.16(f) of \$130 for a non-small entity, must be submitted with the missing items identified in this letter.

### SUMMARY OF FEES DUE:

Total additional fee(s) required for this application is \$1180 for a Large Entity

JUNIPER Exhibit 1003

App. 3, pg. 305

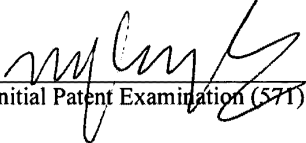
'652 File History 305

- **\$300** Statutory basic filing fee.
- **\$130** Surcharge.
  
- The application search fee has not been paid. Applicant must submit **\$500** to complete the search fee.
- The application examination fee has not been paid. Applicant must submit **\$200** to complete the examination fee for a large entity
  
- Total additional claim fee(s) for this application is **\$50**
  - **\$50** for 1 total claims over 20.

Replies should be mailed to: Mail Stop Missing Parts  
Commissioner for Patents  
P.O. Box 1450  
Alexandria VA 22313-1450

---

*A copy of this notice **MUST** be returned with the reply.*

  
Office of Initial Patent Examination (571) 272-4000, or 1-800-PTO-9199, or 1-800-972-6382

PART 3 - OFFICE COPY

021406  
13328  
U.S. PTO

**UTILITY PATENT APPLICATION TRANSMITTAL**

(New Nonprovisional Applications Under 37 CFR § 1.53(b))

Attorney Docket No.

HAMMP008

**TO THE COMMISSIONER FOR PATENTS:**

Transmitted herewith is the patent application of ( ) application identifier or (X) first named inventor, Ping Pan, entitled PEUDOWIRE PROTECTION, for a(n):

- (X) Original Patent Application.
- ( ) Continuing Application (prior application not abandoned):
  - ( ) Continuation ( ) Divisional ( ) Continuation-in-part (CIP) of prior Application No. \_\_\_\_\_, filed \_\_\_\_\_.
- ( ) Please add after the title of the application "This is a
  - ( ) Continuation ( ) Divisional ( ) Continuation-in-part (CIP) of Application No. \_\_\_\_\_, filed \_\_\_\_\_, which is hereby incorporated by reference."
- (X) This application claims the benefit of U.S. Provisional Application No. 60/653,065 filed February 14, 2005.

112967 U.S. PTO  
11/354569  
021406

Enclosed are:

- (X) Specification; 17 Total Pages. (X) Drawing(s); 7 Total Sheets.
- ( ) Oath or Declaration:
  - ( ) A Newly Executed Combined Declaration and Power of Attorney:
    - ( ) Signed. ( ) Unsigned. ( ) Partially Signed.
  - ( ) A Copy from a Prior Application for Continuation/Divisional (37 CFR § 1.63(d)).
    - ( ) Signed Statement Deleting Inventor(s) Named in the Prior Application. (37 CFR § 163(d)(2)).
- ( ) Power of Attorney. (X) Return Receipt Postcard.
- ( ) Associate Power of Attorney. ( ) A Check in the amount of \$ \_\_\_\_\_ for the Filing Fee.
- ( ) Preliminary Amendment. ( ) Information Disclosure Statement and Form PTO-1449.
- ( ) A Duplicate Copy of this Form for Processing Fee Against Deposit Account.
- ( ) A Certified Copy of Priority Documents (if foreign priority is claimed).
- ( ) Statement(s) of Status as a Small Entity.
- ( ) Statement(s) of Status as a Small Entity Filed in Prior Application, Status Still Proper and Desired.
- (X) Non Publication Request.
- ( ) Other: \_\_\_\_\_

**PLEASE DO NOT CHARGE THE FILING FEE AT THIS TIME.**

Respectfully submitted,

By: *Clover Huang*  
Clover Huang, Reg. No. 55,285

Date: February 14, 2006

Correspondence Address:

**Customer No. 21912**  
Van Pelt, Yi & James LLP  
10050 N. Foothill Blvd.  
Suite 200  
Cupertino, CA 95014  
Telephone: 408-973-2585  
Fax: 408-973-2595

I hereby certify that this is being deposited with the U.S. Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10 on the date indicated below and is addressed to:

Commissioner for Patents  
P. O. Box 1450  
Alexandria, VA 22313-1450

By: *Meghan Long*

Typed Name: Meghan Long

Express Mail Label No.: EV324996756US

Date of Deposit: February 14, 2006

Attorney Docket No. HAMMP008

APPLICATION FOR UNITED STATES PATENT

**PSEUDOWIRE PROTECTION**

By Inventors:

Ping Pan  
San Jose, CA  
A Citizen of the United States of America

Assignee: Hammerhead Systems

VAN PELT, YI & JAMES LLP  
10050 N. Foothill Blvd., Suite 200  
Cupertino, CA 95014  
Telephone (408) 973-2585

JUNIPER Exhibit 1003

App. 3, pg. 308

'652 File History 308

## **PSEUDOWIRE PROTECTION**

### **CROSS REFERENCE TO OTHER APPLICATIONS**

[0001] This application claims priority to U.S. Provisional Patent Application No. 60/653,065 entitled PSEUDO WIRE PROTECTION filed February 14, 2005 which is incorporated herein by reference for all purposes.

### **BACKGROUND OF THE INVENTION**

[0002] In recent years, many networking and telecommunications carriers have deployed Pseudowires to carry Layer-2 (also known as the data link layer of the Open Systems Interconnection (OSI) Reference Model) traffic. A Pseudowire (PW) refers to an emulation of a native service over a network. Examples of the native service include Asynchronous Transfer Mode (ATM), Frame Relay, Ethernet, Time Division Multiplexing (TDM), Synchronous Optical Network (SONET), Synchronous Digital Hierarchy (SDH), etc. Examples of the network include Multiprotocol Label Switching (MPLS), Internet Protocol (IP), etc. More recently, a number of carriers have extended the use of Pseudowires beyond packet encapsulation, and offered Pseudowires as a type of network service. Consequently, data traffic protection and redundancy in environments that use Pseudowire have become critical.

[0003] At the edge of a network, a network edge device such as an edge router may receive multiple Layer-2 flows (also referred to as Attachment Circuits (ACs)). In a typical network supporting Pseudowires, each AC is mapped to a Pseudowire. Ingress packets received mapped to a specific Pseudowire are labeled with an identifier associated with this Pseudowire, and are switched via the Pseudowire. A physical link may support one or more Pseudowires. Ideally, the data flow in a Pseudowire should be protected. In other words, if an active Pseudowire fails, the data flow should be redirected to an alternative Pseudowire to avoid data loss.

**[0004]** Pseudowires can operate over many physical media types. However, existing Pseudowire systems typically provide no protection or very limited protection. For example, there is usually no data protection for Pseudowires on different physical media types, since most network protection schemes, such as APS for SONET, Link Aggregation for Ethernet, do not apply over multiple physical media types.

**[0005]** Some MPLS devices implement schemes such as MPLS Fast Reroute to provide limited data protection. These existing schemes, however, often do not provide adequate protection. Take the following scenario as an example: between two provider edges (PEs), a first tunnel comprising multiple Pseudowires is protected by a second tunnel. Due to network topology constraints, the two tunnels may have different bandwidth. This is a possible scenario in an MPLS Fast Reroute operation. In this example, the second tunnel may have lower bandwidth than that of the first one. If the first tunnel should fail, the amount of data that needs to be redirected through the second tunnel may exceed the capacity of the second tunnel. Furthermore, existing protocols typically do not provide a way of determining which data gets priority. Thus, certain mission critical data may be dropped while other less critical data may pass through.

**[0006]** It would be desirable to have a way to provide better Pseudowire protection and to have more control during switchover. It would also be desirable if the protection scheme could be implemented without significant changes to existing protocols and devices.



## **BRIEF DESCRIPTION OF THE DRAWINGS**

**[0007]** Various embodiments of the invention are disclosed in the following detailed description and the accompanying drawings.

**[0008]** FIGS. 1A and 1B are block diagrams illustrating an embodiment of a single-hop Pseudowire system and an embodiment of a multi-hop Pseudowire system, respectively.

**[0009]** FIG. 2 is a flowchart illustrating an embodiment of a process of providing data protection using Pseudowires.

**[0010]** FIG. 3A is a flowchart illustrating another embodiment of a process of providing data protection using Pseudowires.

**[0011]** FIG. 3B is a flowchart illustrating how the Pseudowire is used, according to some embodiments.

**[0012]** FIG. 4 is a data structure diagram illustrating an embodiment of a Pseudowire protection configuration parameter that specifies several protection-related properties of the Pseudowire.

**[0013]** FIG. 5 is a flowchart illustrating an example process of using the priorities during switchover.

**[0014]** FIG. 6 is a diagram illustrating an example in which preemption takes place during a switchover operation.

## **DETAILED DESCRIPTION**

**[0015]** The invention can be implemented in numerous ways, including as a process, an apparatus, a system, a composition of matter, a computer readable medium such as a computer readable storage medium or a computer network wherein program instructions are sent over optical or electronic communication links. In this specification, these implementations, or any other form that the invention may take, may be referred to as techniques. A component such as a processor or a memory described as being configured to perform a task includes both a general component that is temporarily configured to perform the task at a given time or a specific component that is manufactured to perform the task. In general, the order of the steps of disclosed processes may be altered within the scope of the invention.

**[0016]** A detailed description of one or more embodiments of the invention is provided below along with accompanying figures that illustrate the principles of the invention. The invention is described in connection with such embodiments, but the invention is not limited to any embodiment. The scope of the invention is limited only by the claims and the invention encompasses numerous alternatives, modifications and equivalents. Numerous specific details are set forth in the following description in order to provide a thorough understanding of the invention. These details are provided for the purpose of example and the invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the invention is not unnecessarily obscured.

**[0017]** Providing protection to network traffic using one or more Pseudowires is disclosed. In some embodiments, a Pseudowire protection configuration parameter is sent to a destination node. A Pseudowire configuration acknowledgment from the destination node is received. If a Pseudowire is allowed to be established according to the Pseudowire configuration acknowledgment, it is established based at least in part on the Pseudowire protection configuration parameter. In embodiments where the

Pseudowire is established as a standby Pseudowire configured to protect one or more primary Pseudowires, in the event that a primary Pseudowire fails to transfer network traffic for reasons such as network congestion, equipment failure, etc., network traffic that is originally designated to be transferred on the primary Pseudowire(s) is switched from the primary Pseudowire(s) to the standby Pseudowire.

**[0018]** The protection technique is applicable to both single-hop and multi-hop systems. FIGS. 1A and 1B are block diagrams illustrating an embodiment of a single-hop Pseudowire system and an embodiment of a multi-hop Pseudowire system, respectively. Configuring and switching the Pseudowire will be discussed in more detail below.

**[0019]** In the example shown in FIG. 1A, system 100 is a single-hop system where the nodes in the system all belong to the same carrier network. Within each carrier network, all network nodes and facility are under a common administrative control. A service provider company may own multiple carrier networks in different regions. As used herein, a node refers to a networked device. In this case, the nodes in the system are provider edges (PEs) A, B, C, and D, which all belong to the same carrier network. Ingress data received by attachment circuits 112 of PE A designated for PE B may be sent via a label switched path (LSP) through PEs A, C, and B, or an LSP through PEs A, D, and B. The first LSP comprises Pseudowires 102, 104 and 106, and the second LSP comprises Pseudowires 108 and 110. In this example, the Pseudowire connections between PEs are established using the Label Distribution Protocol (LDP). The connections are based on LDP sessions. Each LDP session is to connect two local or remote nodes. There may be multiple paths interconnecting any two nodes in the network. Thus, for each LDP session, there may be multiple LDP Hello Adjacencies, one LDP Hello Adjacency per path. For purposes of example, throughout this specification, LDP is used as the communication protocol between nodes. Other appropriate protocols may also be used.

**[0020]** In the example shown in FIG. 1B, system 150 is a multi-hop system since

it includes multiple carrier networks. Carrier networks 1-6 form autonomous systems 1-6, respectively. Each autonomous system includes one or more networks that are controlled by a carrier. For purposes of illustration, three Pseudowires are shown in this example to transfer data between PE 1A and PE 3B: a first Pseudowire comprising a path via autonomous systems 1, 2, and 3, a second Pseudowire comprising a path via autonomous systems 1, 6, and 3, and a third Pseudowire comprising a path via autonomous systems 1, 4, 5, and 3. Other Pseudowire formations are possible. At the source node PE 1A, data packets to be sent via a particular Pseudowire are labeled with an identifier associated with the Pseudowire, forwarded on to the next provider edge on one Pseudowire segment, and forwarded again if necessary until the packets reach the destination node 3B.

**[0021]** FIG. 2 is a flowchart illustrating an embodiment of a process of providing data protection using Pseudowires. Process 200 may be implemented on a source node such as A or 1A of systems 100 and 150, or on an independent management agent that communicates with the source node. For purposes of illustration, the process is shown as implemented on a source node in the following example. The process initializes when a connection session is established between the source node and the destination node (202). A Pseudowire protection configuration parameter for configuring a Pseudowire based on the connection session is sent (204). The Pseudowire protection configuration parameter includes one or more fields that specify certain protection properties associated with the Pseudowire. It may be sent to the destination node or a management agent that communicates with the destination node. Details of the configuration parameter will be discussed further below.

**[0022]** Once the destination node (or its associated management agent) receives the Pseudowire protection configuration parameter, it determines whether it will accept the Pseudowire protection configuration and allow a standby Pseudowire to be established. Depending on implementation, the destination node determines whether to accept the protection configuration based on factors such as traffic condition, number of existing Pseudowires, priority information, etc. The destination node may reject the

protection request for a number of reasons. For example, the destination node does not support Pseudowire protection mechanism as described here. If a standby Pseudowire may be established, the destination node accepts it and configures the Pseudowire based at least in part on the configuration parameters. In some embodiments, the destination node adds the Pseudowire to a table of Pseudowires. A corresponding Pseudowire configuration acknowledgment is generated, indicating whether the destination node has accepted the Pseudowire configuration. The Pseudowire configuration acknowledgment is sent to the source node. In some embodiments, as a part of the LDP process, a MPLS label for the data packets traversing through the standby Pseudowire is assigned.

**[0023]** At the source node, once the Pseudowire configuration acknowledgment is received (206), it is examined to determine whether the Pseudowire configuration has been accepted (208). If, according to the Pseudowire configuration acknowledgment, the Pseudowire configuration has been accepted by the destination, a standby Pseudowire is established based at least in part on the Pseudowire protection configuration parameter and may be used as such (210). If, however, the Pseudowire configuration has not been accepted, the process performs appropriate exception handling, such as re-sending the Pseudowire protection configuration parameter (212).

**[0024]** FIG. 3A is a flowchart illustrating another embodiment of a process of providing data protection using Pseudowires. Process 300 may be implemented on a PE, on an independent management agent, or the like. For purposes of illustration, in the following example, the process is initiated and carried out on a PE source node.

**[0025]** Process 300 begins with the initialization of an LDP session (302). According to the negotiation scheme based on LDP, the source node exchanges messages with the destination node and establishes an LDP Hello Adjacency (304). A Pseudowire setup request that includes a Pseudowire protection configuration parameter is sent to the destination node (or its associated management agent), requesting that a standby Pseudowire be established over the LDP Hello Adjacency (306). In some embodiments, multiple LDP Hello Adjacencies are available for Pseudowire setup, thus multiple setup

requests are sent, and the destination node processes the requests and maps Pseudowires to appropriate LDP Hello Adjacencies. In some embodiments, the source node dynamically determines which LDP Hello Adjacency among the available connections is to be configured as a standby Pseudowire, and directs its setup request accordingly. The dynamic determination may be based on, among other things, bandwidth availability on the adjacency path.

**[0026]** In some embodiments, the request is sent as a LDP Label Mapping Message. The configuration parameter is used to configure various properties of the Pseudowire, including protection type, protection scheme, priority, etc. Further details of the configuration parameters are discussed below. In some embodiments, multiple LDP Hello Adjacencies are established and the source node sends multiple Pseudowire setup requests to configure Pseudowires over these LDP Hello Adjacencies.

**[0027]** In this example, upon receiving a Pseudowire setup request, the destination node maps the request to the appropriate LDP Hello Adjacency. If the mapping is successful, the Pseudowire is established. Sometimes, however, the mapping and consequently the Pseudowire setup may fail for reasons such as network congestion, resource limitation, equipment failure, etc. The destination node sends a Pseudowire configuration acknowledgment to the source node. In this example, the Pseudowire configuration acknowledgment is an LDP acknowledgement indicating whether a particular Pseudowire has been successfully established. Once the source node receives the acknowledgement (308), it determines whether the configuration has been accepted by the destination (310). If the configuration has been accepted, a standby Pseudowire is successfully established based at least in part on the Pseudowire protection configuration parameter, and the source and destination nodes can start using the standby Pseudowire to protect other Pseudowires (312). If, however, the acknowledgment indicates that the configuration has not been accepted and a Pseudowire has not been successfully established, appropriate exception handling measures such as resending the Pseudowire protection configuration parameter are taken (314).

**[0028]** Process 300 is applicable to both single-hop and multi-hop systems. In a single-hop system, the source node and the destination node correspond to a source PE and a destination PE on the network and the process is used to configure a standby Pseudowire between the PEs. In a multi-hop system, the process may be repeated by the PEs on various carrier networks to establish Pseudowire segments. For example, in system 150 of FIG. 1B, PE 1A can use process 300 to establish a Pseudowire segment with PE 6A, and PE 6A can use the same process to establish a Pseudowire segment with PE 6B, which can use the same process to establish a Pseudowire segment with PE 3B.

**[0029]** FIG. 3B is a flowchart illustrating how the Pseudowire is used, according to some embodiments. Process 350 may be implemented on the source node, the destination node, or both. In this example, the designation of the Pseudowire is first determined (352). The designation may be configured by a system administrator, in a Pseudowire configuration process, or any other appropriate means. If the Pseudowire is designated as a primary Pseudowire, it is configured to carry network traffic (354). In the event that a primary Pseudowire fails (356), the nodes associated with the Pseudowire will attempt to switch the traffic over to the standby Pseudowire by sending a switchover request to the Pseudowire (358). As will be shown in more detail below, in some embodiments, whether the traffic on the primary Pseudowire can preempt the traffic on the standby Pseudowire and be switched over depends on priority configuration of the Pseudowires.

**[0030]** If it is designated as a standby Pseudowire, it enters into standby mode to provide protection to one or more primary Pseudowires (360). In some embodiments, the standby Pseudowire carries network traffic during normal operation. It is ready to take over traffic from the primary Pseudowire if necessary. If a switchover request is received from a primary Pseudowire (362), traffic on the primary Pseudowire is switched over to the standby Pseudowire. In some embodiments, the switchover only occurs if the priority comparison of the primary and standby Pseudowires indicates the switchover is allowed.

**[0031]** Optionally, during the operation, if a Pseudowire is no longer needed, the source node can send a withdraw request over the Pseudowire and the destination node disassociates the Pseudowire with the LDP Hello Adjacency to break the Pseudowire connection.

**[0032]** FIG. 4 is a data structure diagram illustrating an embodiment of a Pseudowire protection configuration parameter that specifies several protection-related properties of the Pseudowire. In this example, Pseudowire protection configuration parameter 400 includes four fields: protection scheme, protection type, domain type, and priority. A field may have one or more subfields. For example, the priority field is shown to include a holding priority and a setup priority. One or more of the fields and/or subfields may be used in various embodiments. Other appropriate fields may also be implemented. In the example shown, the fields are numerical values that map to appropriate property values.

**[0033]** In some embodiments, one of the following Pseudowire protection schemes is used to set up the Pseudowires: 1+1, 1:1, 1:N or M:N. The protection scheme field is used to indicate which protection scheme is used in the system setup. A specific protection scheme corresponds to a field value. For example, 1+1 maps to 0, 1:1 maps to 1, and so on. In a system implementing a 1+1 protection scheme, the same traffic is sent over two parallel Pseudowires and the receiver selects one traffic stream at a time. In a system implementing a 1:1 protection scheme, one Pseudowire is used to protect another Pseudowire. Similarly, in a 1:N system (e.g. MPLS Facility Backup), one Pseudowire is used to protect N other Pseudowires, and in a M:N system M Pseudowires are used to protect N other Pseudowires.

**[0034]** The protection type field is used to configure the standby mode of the Pseudowire. In some embodiments, cold, warm, and hot standby modes are supported. Other appropriate standby modes may be implemented in other embodiments. In some embodiments, in cold standby mode configuration, once network failure on a Pseudowire carrying network traffic is detected, a standby Pseudowire is selected from the remaining



functional Pseudowires, and traffic is redirected to the standby Pseudowire. In some embodiments with warm standby mode configuration, one or more standby Pseudowires are established before any network failure has occurred. These standby Pseudowires, however, are not maintained or used to transport data until a network failure is detected. Upon failure detection, the source or destination nodes will modify the data-plane and switch data traffic over to the standby Pseudowire(s). In some embodiments with hot standby mode configuration, one or more standby Pseudowires are pre-established and maintained at both control-plane and data-plane, so that once a network failure is detected, data traffic is directly switched over to the standby Pseudowire(s).

**[0035]** The domain type field indicates whether the Pseudowire is configured in a single-hop environment where all the nodes of the Pseudowire belong to the same carrier network, or a multi-hop environment where the Pseudowire includes nodes on several carrier networks. This is because the intermediate may process single-hop and multi-hop Pseudowire differently.

**[0036]** The priority field indicates the preference level of a Pseudowire in preempting other Pseudowires during switchover. In the event of a network failure, the edge nodes will preferentially provide protection according to the priority setting of the Pseudowires. In a situation where network resources (such as bandwidth) are limited, data sent on a higher priority Pseudowire is more likely to be protected than data sent on a lower priority Pseudowire. In some embodiments, the priority field includes two subfields: a holding priority and a setup priority. The holding priority indicates the relative priority of a currently active Pseudowire with respect to other Pseudowires when the latter attempt to preempt the former's use of the data link. Stated another way, it determines how easily a currently active Pseudowire gives up its hold on a data link upon request. The setup priority indicates the relative priority of a Pseudowire during the setup process.

**[0037]** FIG. 5 is a flowchart illustrating an example process of using the priorities during switchover. Process 500 may be implemented on an edge node, an independent

management agent, or the like. In this example, process 500 initiates when a network failure has been detected (502). It is determined whether preemption is required (504). Preemption is required when the failed link carries more Pseudowire traffic than the available bandwidth on the standby link. If preemption is not required, the Pseudowire(s) may directly switchover (506). If, however, preemption is required, the setup priorities of the Pseudowires on the failed link are compared and the Pseudowire with the highest setup priority is selected (508). The setup priority of the selected Pseudowire is compared to the holding priority of the standby Pseudowire (510). If the setup priority is greater than the holding priority, traffic on the selected Pseudowire is switched over to the standby Pseudowire (506). If, however, the setup priority is no greater than the holding priority, no switchover takes place and the standby Pseudowire continues to transfer its own data and the data on the failed Pseudowires is lost (514).

**[0038]** FIG. 6 is a diagram illustrating an example in which preemption takes place during a switchover operation. In this example, Pseudowires 600, 602 and 604 are active, primary Pseudowires carrying traffic. Pseudowire 604 is used as the standby Pseudowire. Pseudowire 600 has a holding priority and a setup priority of 10 and 11, respectively, Pseudowire 602 has priorities of 11 and 12, and Pseudowire 604 has priorities of 9 and 9. Thus, if the link on which Pseudowires 600 and 602 operate fails, the nodes will initiate switchover using Pseudowire 604. A comparison of the setup priority of Pseudowires 600 and 602 indicates that Pseudowire 602 has a higher setup priority, thus 602 is given preference in the switchover. The setup priority of Pseudowire 602 is compared with the holding priority of Pseudowire 604. Since 602's setup priority is greater than 604's holding priority, data on 602 preempts data on 604 and takes over the link.

**[0039]** Providing protection to network traffic using one or more Pseudowires has been disclosed. Pseudowire protection improves the reliability of Pseudowire services. Pseudowires are better controlled by appropriately configuring the properties of Pseudowires and without requiring significant changes to existing protocols and devices.

**[0040]** Although the foregoing embodiments have been described in some detail for purposes of clarity of understanding, the invention is not limited to the details provided. There are many alternative ways of implementing the invention. The disclosed embodiments are illustrative and not restrictive.

**[0041]** WHAT IS CLAIMED IS:

## CLAIMS

1. A method of providing protection to network traffic, comprising:
  - sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node;
  - 5 receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and
  - in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire;
  - 10 wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter.
2. A method as recited in Claim 1, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.
3. A method as recited in Claim 1, wherein the standby Pseudowire is configured to  
15 provide protection to at least one primary Pseudowire, and in the event that the primary Pseudowire fails to transfer network traffic, switching network traffic from at least one of said at least one primary Pseudowire to the standby Pseudowire.
4. A method as recited in Claim 1, wherein the standby Pseudowire is dynamically selected from a plurality of connections.
- 20 5. A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a domain type.
6. A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a protection type.
7. A method as recited in Claim 1, wherein the Pseudowire protection configuration  
25 parameter includes a protection scheme.
8. A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter includes a priority.
9. A method as recited in Claim 1, further including determining whether to preempt

existing traffic on the standby Pseudowire, the determination being based at least in part on a priority associated with the standby Pseudowire.

10. A method as recited in Claim 1, wherein the Pseudowire protection configuration parameter is established using the Label Distribution Protocol (LDP).

5 11. A system for providing protection to network traffic, comprising:  
a processor configured to:  
send a Pseudowire protection configuration parameter for configuring a  
standby Pseudowire between a source node and a destination node; and  
receive a Pseudowire configuration acknowledgement indicating whether  
10 the Pseudowire protection configuration parameter has been accepted by the  
destination node; and  
in the event that the Pseudowire protection configuration parameter has  
been accepted by the destination node, use the standby Pseudowire;  
wherein the standby Pseudowire is configured based at least in part on the  
15 Pseudowire protection configuration parameter; and  
a memory coupled to the processor, configured to provide the processor with  
instructions.

12. A system as recited in Claim 11, wherein the standby Pseudowire is configured to provide protection to at least one primary Pseudowire.

20 13. A system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a domain type.

14. A system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a protection type.

25 15. A system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a protection scheme.

16. A system as recited in Claim 11, wherein the Pseudowire protection configuration parameter includes a priority.

17. A computer program product for configuring a Pseudowire between a source node

and a destination node, the computer program product being embodied in a computer readable medium and comprising computer instructions for:

    sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node;

5           receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node; and

    in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire;

10           wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter.

18.    A computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a domain type.

19.    A computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a protection type.

20.    A computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a protection scheme.

21.    A computer program product as recited in claim 17, wherein the Pseudowire protection configuration parameter includes a priority.

## **PSEUDOWIRE PROTECTION**

### **ABSTRACT OF THE DISCLOSURE**

Providing protection to network traffic includes sending a Pseudowire protection configuration parameter for configuring a standby Pseudowire between a source node and a destination node, receiving a Pseudowire configuration acknowledgement indicating whether the Pseudowire protection configuration parameter has been accepted by the destination node, and in the event that the Pseudowire protection configuration parameter has been accepted by the destination node, using the standby Pseudowire, wherein the standby Pseudowire is configured based at least in part on the Pseudowire protection configuration parameter.

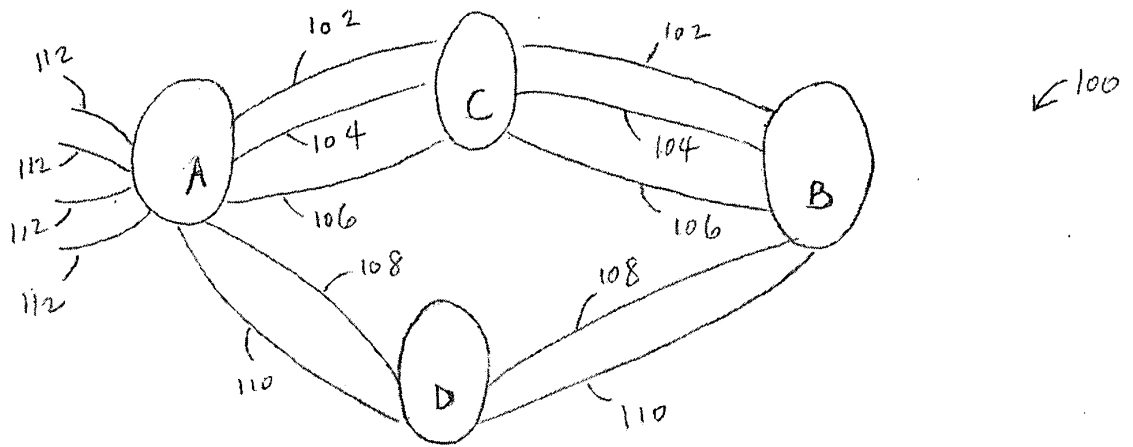


FIG. 1A

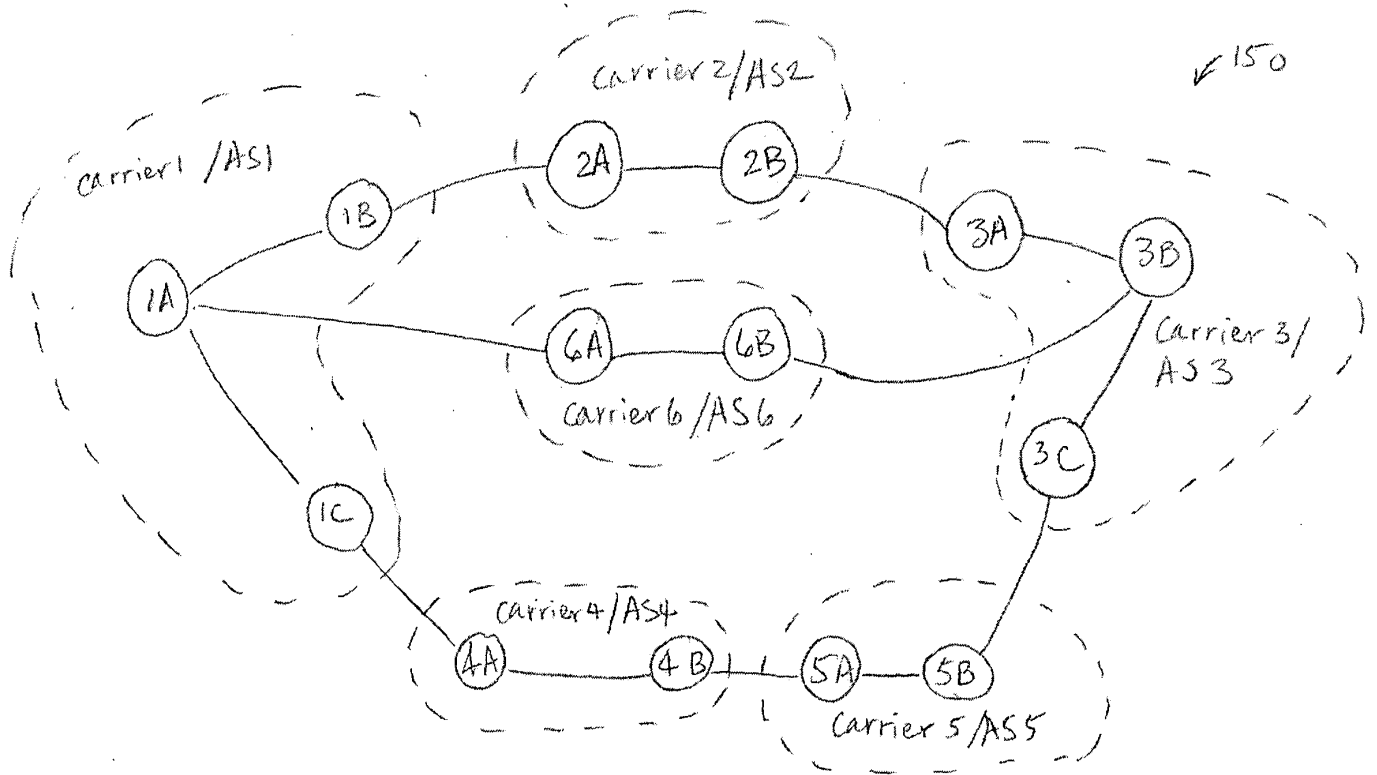


FIG. 1B

**BEST AVAILABLE COPY**

JUNIPER Exhibit 1003

App. 3, pg. 326

'652 File History 326



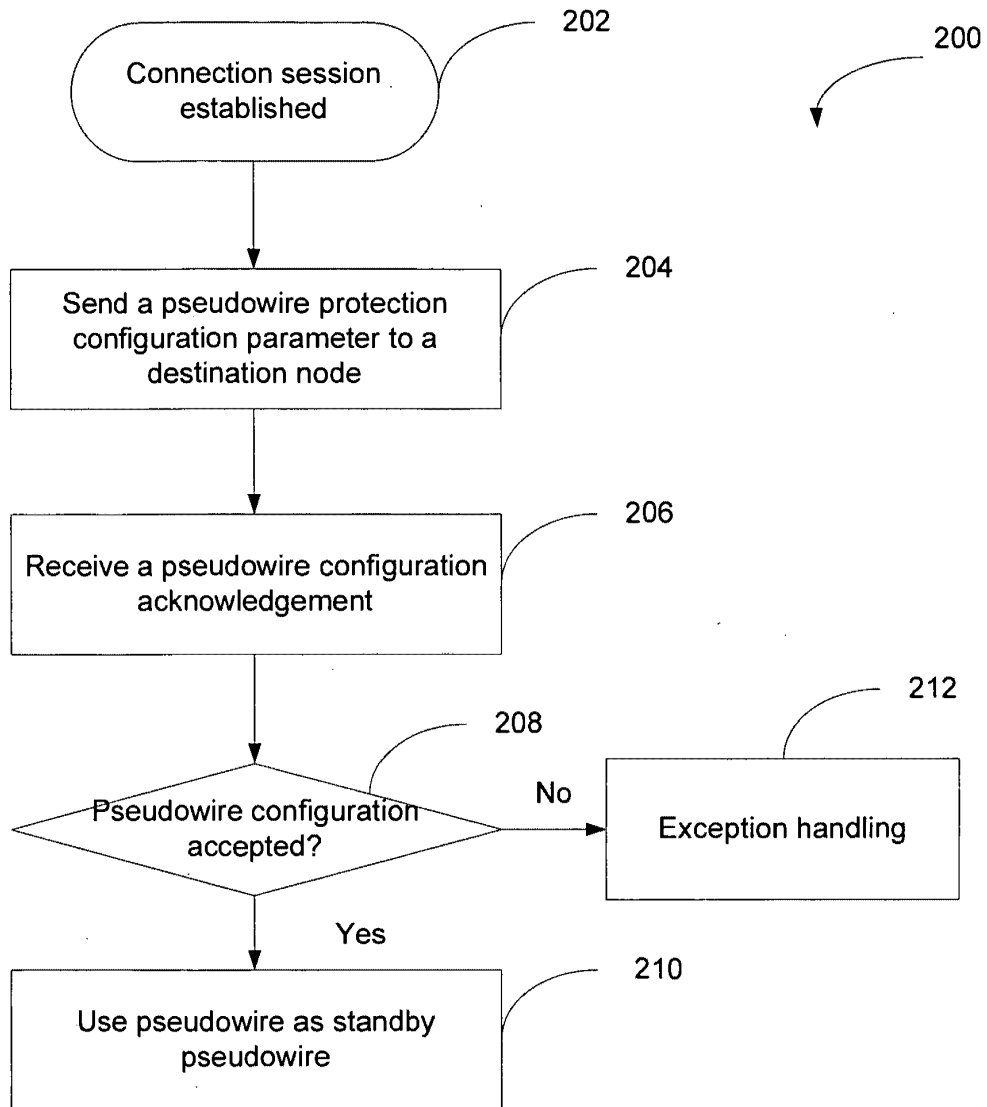


FIG. 2

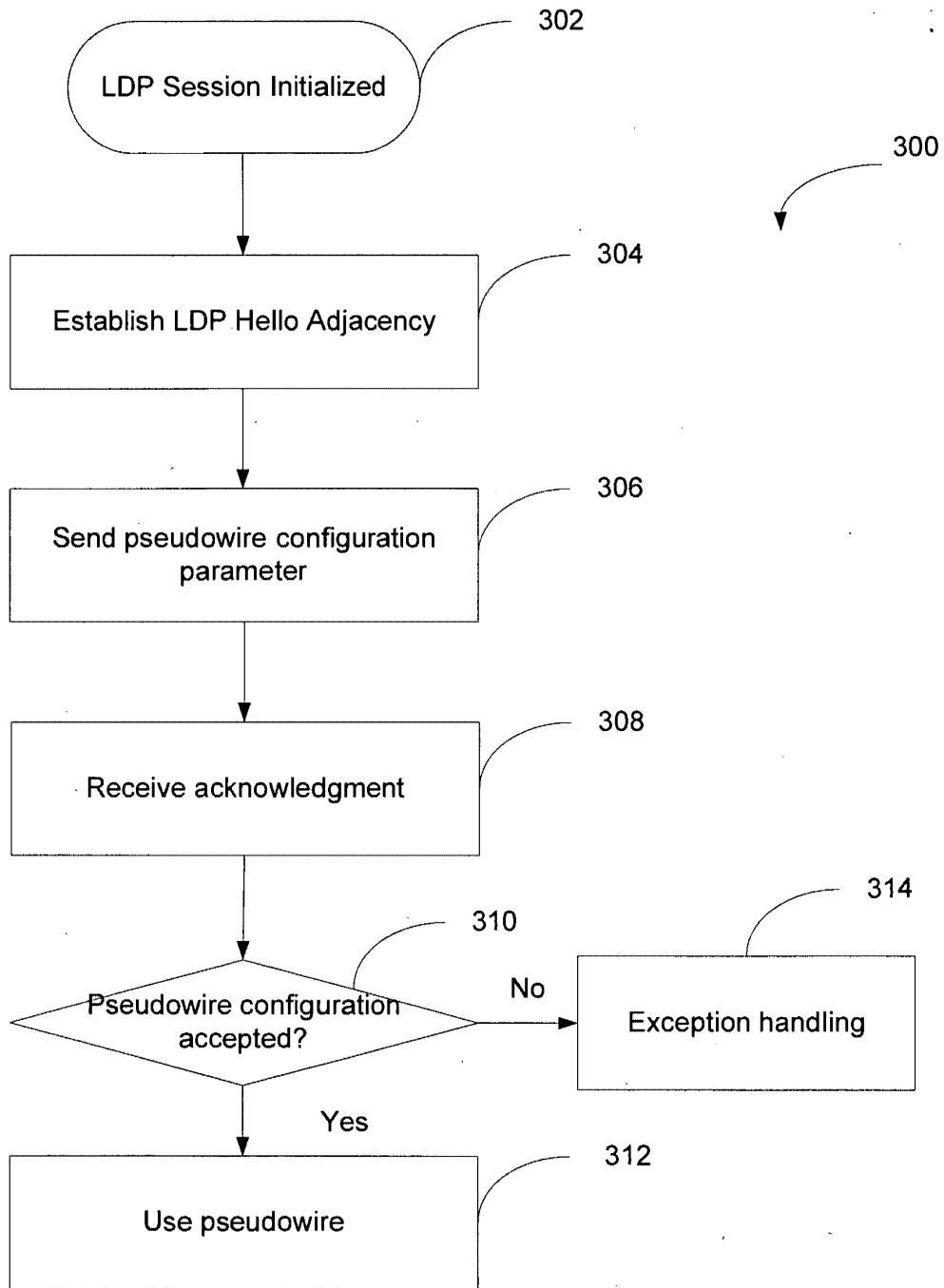


FIG. 3A

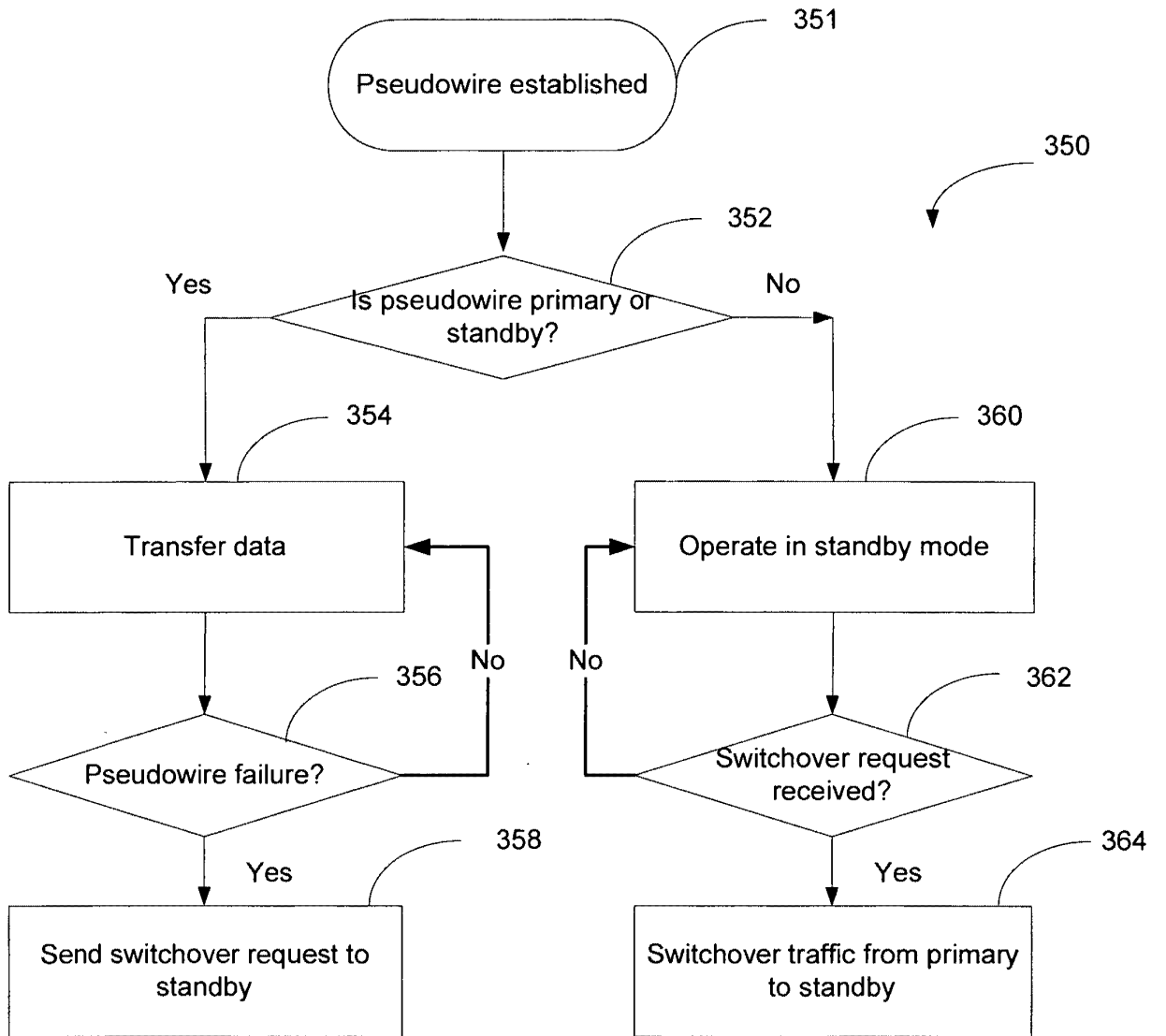


FIG. 3B

protection scheme	protection type	domain type	priority	
01	11	0	10	12

↗  
400

FIG. 4

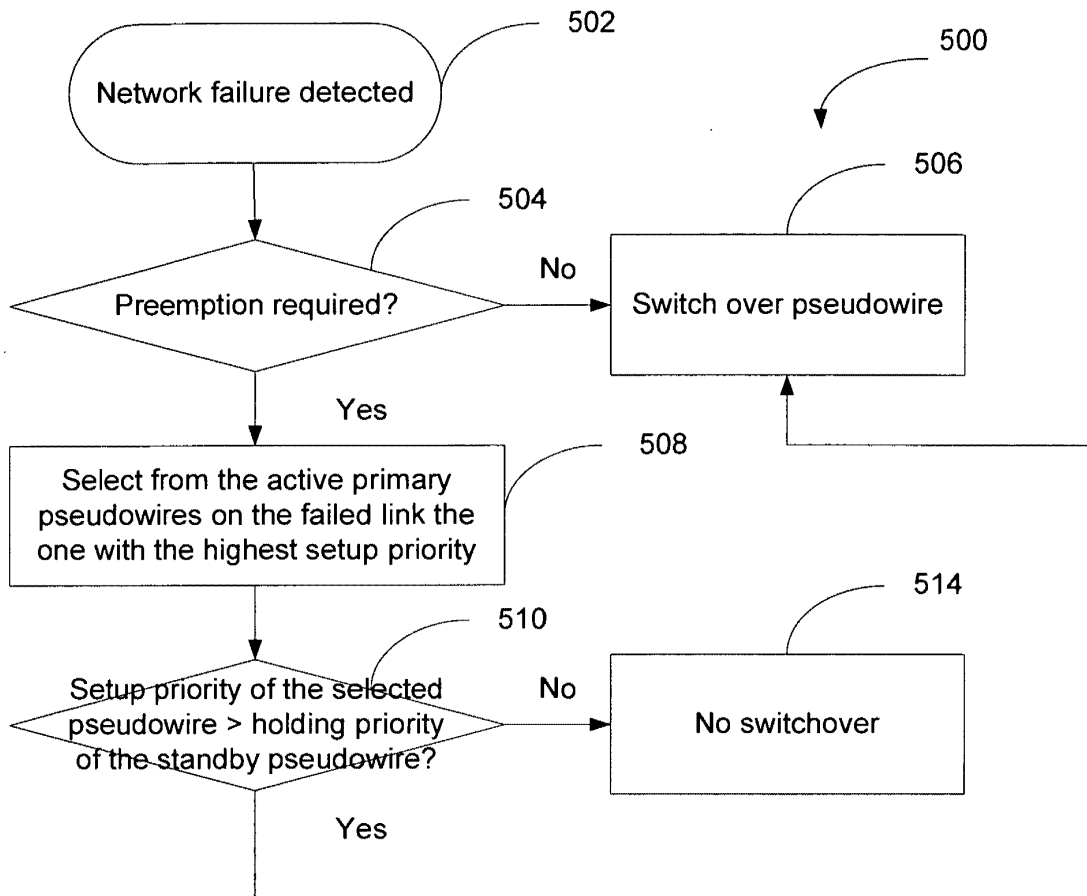


FIG. 5

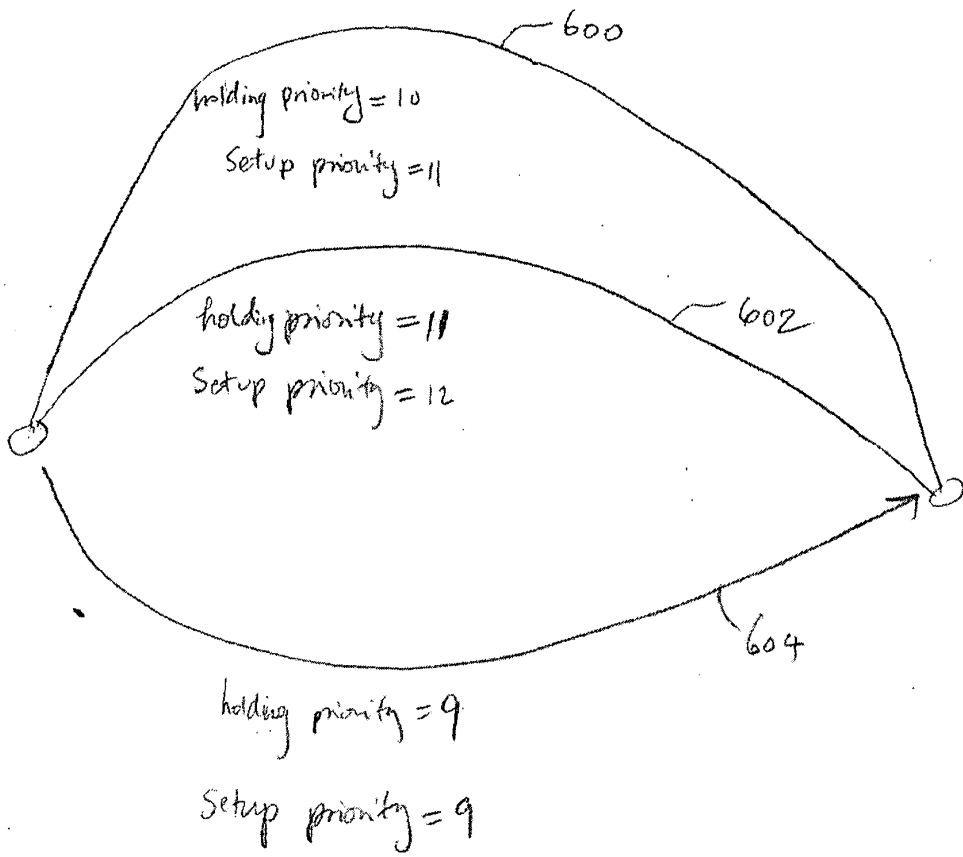


FIG. 6

BEST AVAILABLE COPY

JUNIPER Exhibit 1003

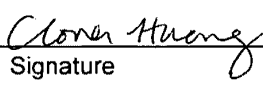
App. 3, pg. 332

'652 File History 332

<b>NONPUBLICATION REQUEST UNDER 35 U.S.C. 122(b)(2)(B)(i)</b>	First Named Inventor		Ping Pan
	Title	PSEUDOWIRE PROTECTION	
	Attorney Docket Number		HAMMP008

I hereby certify that the invention disclosed in the attached application **has not and will not be** the subject of an application filed in another country, or under a multilateral agreement, that requires publication at eighteen months after filing.

I hereby request that the attached application not be published under 35 U.S.C. 122(b).

 _____ Signature	2/14/2006 _____ Date
Clover Huang _____ Typed or printed name	55,285 _____ Registration Number, if applicable
408-973-2594 _____ Telephone Number	

This request must be signed in compliance with 37 CFR 1.33(b) and submitted with the application **upon filing**.

Applicant may rescind this nonpublication request at any time. If applicant rescinds a request that an application not be published under 35 U.S.C. 122(b), the application will be scheduled for publication at eighteen months from the earliest claimed filing date for which a benefit is claimed.

If applicant subsequently files an application directed to the invention disclosed in the attached application in another country, or under a multilateral international agreement, that requires publication of applications eighteen months after filing, the applicant **must** notify the United States Patent and Trademark Office of such filing within forty-five (45) days after the date of the filing of such foreign or international application. **Failure to do so will result in abandonment of this application (35 U.S.C. 122(b)(2)(B)(iii)).**

This collection of information is required by 37 CFR 1.213(a). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 6 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

<b>PATENT APPLICATION FEE DETERMINATION RECORD</b> Substitute for Form PTO-875	Application or Docket Number <b>11,354,569</b>
---	---

APPLICATION AS FILED – PART I			SMALL ENTITY		OTHER THAN SMALL ENTITY	
	(Column 1)	(Column 2)				
FOR	NUMBER FILED	NUMBER EXTRA	RATE (\$)	FEE (\$)	RATE (\$)	FEE (\$)
BASIC FEE (37 CFR 1.16(a), (b), or (c))						<b>300</b>
SEARCH FEE (37 CFR 1.16(k), (l), or (m))						<b>500</b>
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))						<b>200</b>
TOTAL CLAIMS (37 CFR 1.16(i))	<b>21</b>	minus 20 = <b>1</b>	X\$ 25		X\$50	<b>50</b>
INDEPENDENT CLAIMS (37 CFR 1.16(h))	<b>3</b>	minus 3 = *	X\$100		X\$200	
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is \$250 (\$125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR					
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))			180		360	
			TOTAL		TOTAL	<b>1050</b>

\* If the difference in column 1 is less than zero, enter "0" in column 2.

APPLICATION AS AMENDED – PART II					SMALL ENTITY		OTHER THAN SMALL ENTITY			
		(Column 1)	(Column 2)	(Column 3)						
<b>AMENDMENT A</b>		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE (\$)	ADDITIONAL FEE (\$)	RATE (\$)	ADDITIONAL FEE (\$)	
	Total (37 CFR 1.16(i))	*	Minus	**	=	X =		X =		
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X =		X =		
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))					180		360		
					TOTAL ADD'T FEE		TOTAL ADD'T FEE			

		(Column 1)	(Column 2)	(Column 3)						
<b>AMENDMENT B</b>		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE (\$)	ADDITIONAL FEE (\$)	RATE (\$)	ADDITIONAL FEE (\$)	
	Total (37 CFR 1.16(i))	*	Minus	**	=	X =		X =		
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X =		X =		
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))					N/A		N/A		
					TOTAL ADD'T FEE		TOTAL ADD'T FEE			

\* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.  
 \*\* If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".  
 \*\*\* If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".  
 The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.





(19) **United States**

(12) **Patent Application Publication**  
**Hofmeister et al.**

(10) **Pub. No.: US 2004/0156313 A1**

(43) **Pub. Date: Aug. 12, 2004**

(54) **METHOD AND APPARATUS FOR PERFORMING DATA FLOW INGRESS/EGRESS ADMISSION CONTROL IN A PROVIDER NETWORK**

(52) **U.S. Cl. .... 370/229**

(57) **ABSTRACT**

(76) **Inventors: Ralph Theodore Hofmeister, Los Altos, CA (US); Ping Pan, San Jose, CA (US)**

**Correspondence Address:**  
**BIRCH STEWART KOLASCH & BIRCH**  
**PO BOX 747**  
**FALLS CHURCH, VA 22040-0747 (US)**

A method, apparatus and network for transporting layer-2 frames, such as Ethernet MAC, ATM AAL5, and Frame Relay, over MPLS, SONET/SDH, or OTN optical transport networks as well as electrical transport networks is disclosed. The method establishes "pseudo-wires" between, for example, routers, Layer-2 packet switches, or SONET/SDH switches. Inter-related ingress and egress resource tables may be used by provider edge nodes to negotiate consistently managed data tunnels across a provider network on behalf of data flowing from/to a diverse base of customer edge nodes. Detailed network resource information particular to each of the data flows is exchanged between provider edge nodes during the creation of pseudo-wires. Admission control algorithms are applied at the ingress and egress points in order to manage the data flows into a provider network and exiting from a provider network to customer equipment. By applying pseudo-wire shuffling and preemption techniques, the providers can make better use of their network resources by admitting more pseudo-wires.

(21) **Appl. No.: 10/769,891**

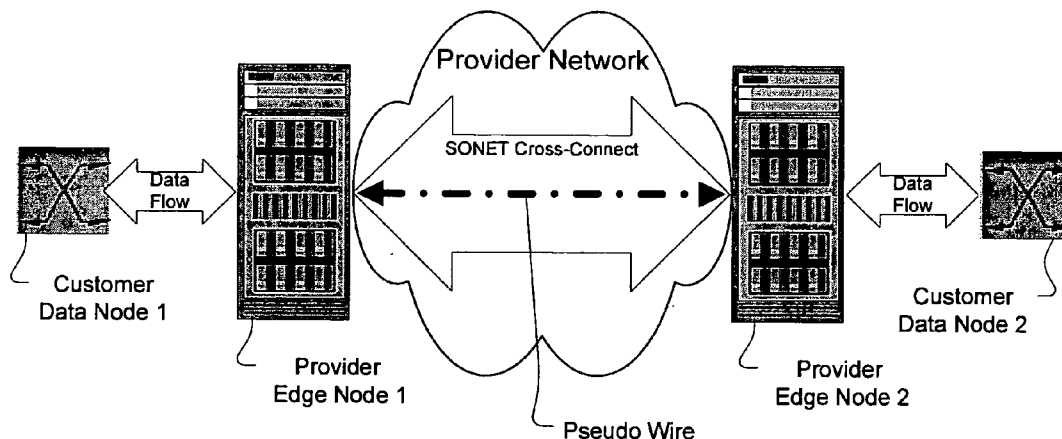
(22) **Filed: Feb. 3, 2004**

**Related U.S. Application Data**

(60) **Provisional application No. 60/444,456, filed on Feb. 3, 2003. Provisional application No. 60/444,440, filed on Feb. 3, 2003.**

**Publication Classification**

(51) **Int. Cl.<sup>7</sup> ..... H04L 12/26**



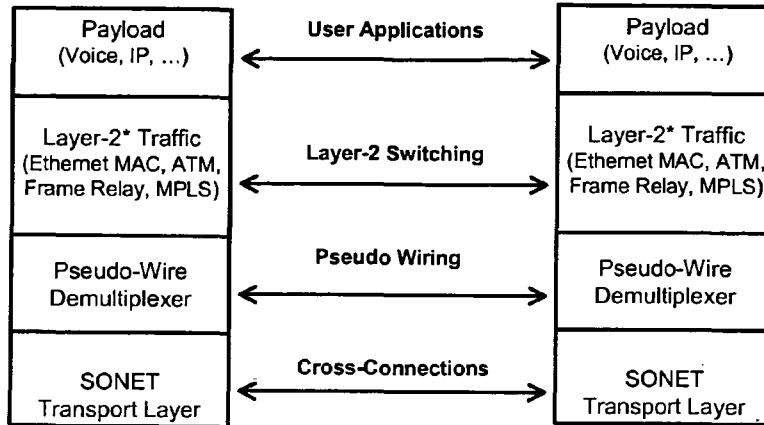


Figure 1

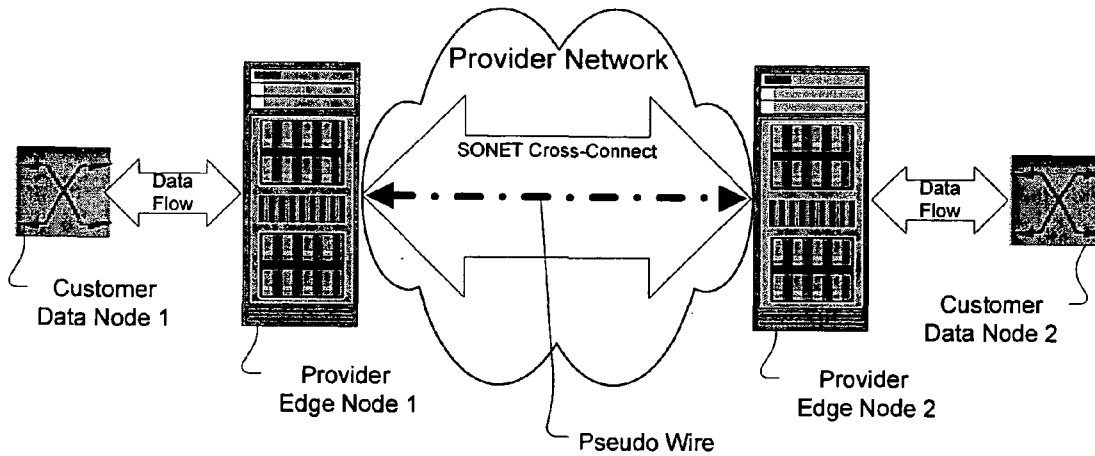


Figure 2

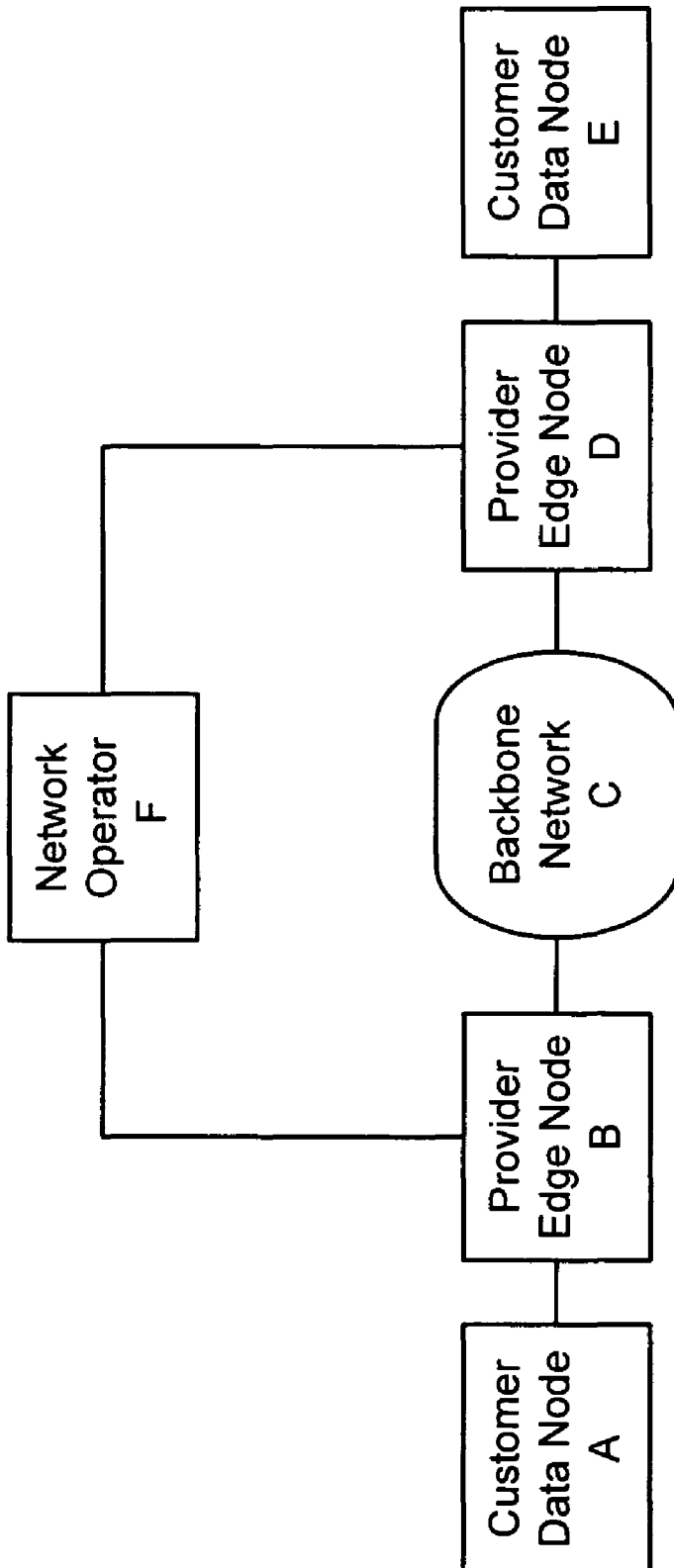


Figure 3

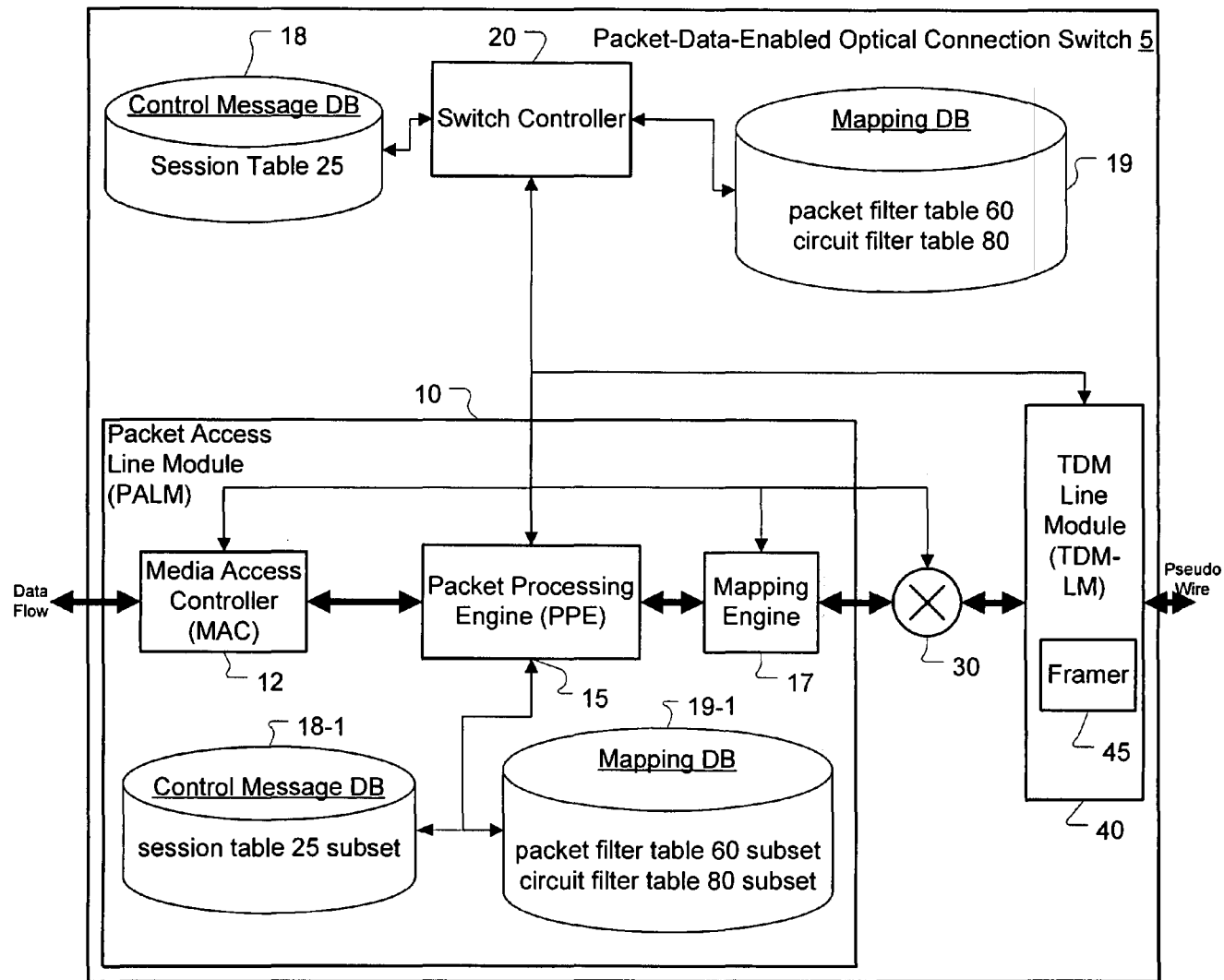


Figure 4

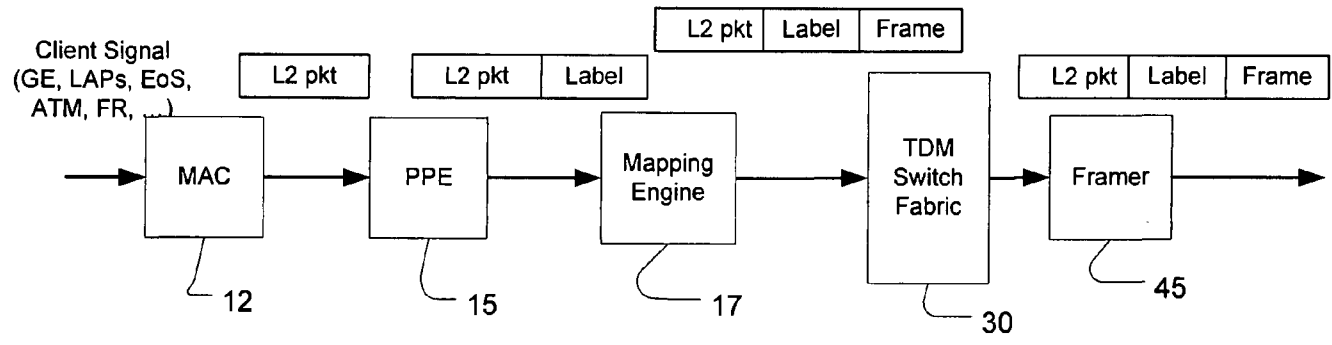


Figure 5

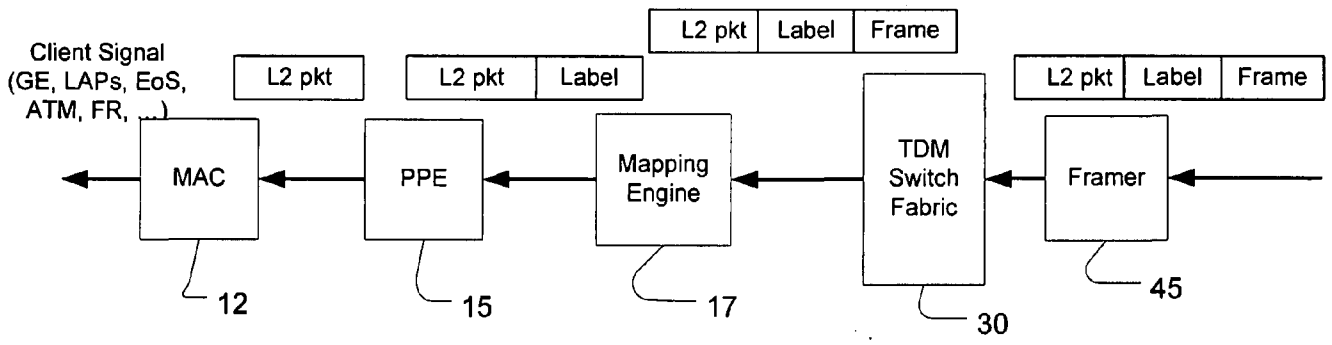


Figure 6

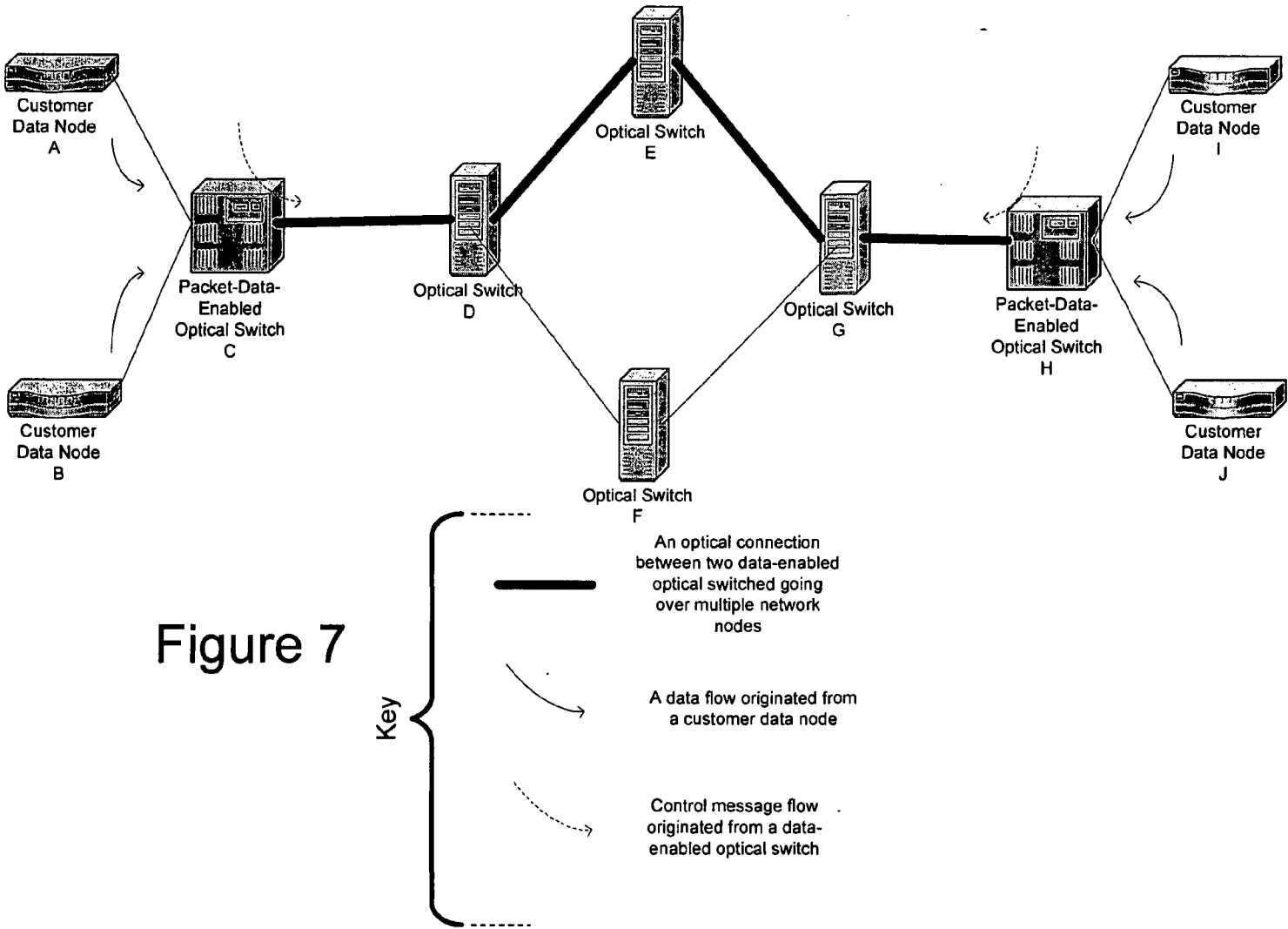


Figure 7

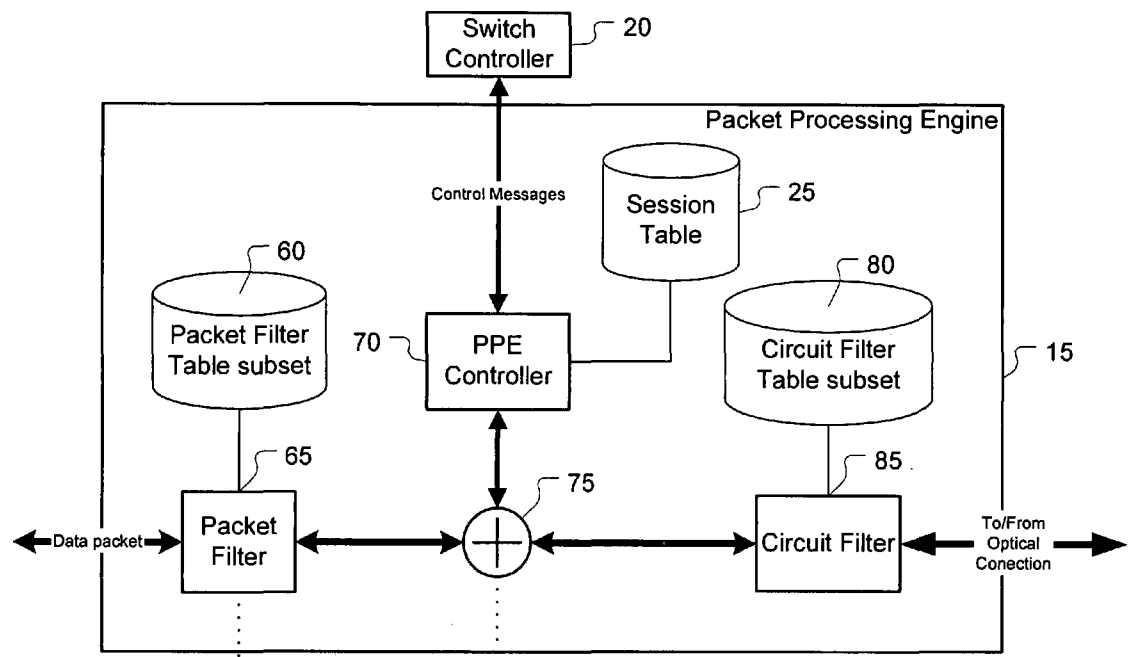
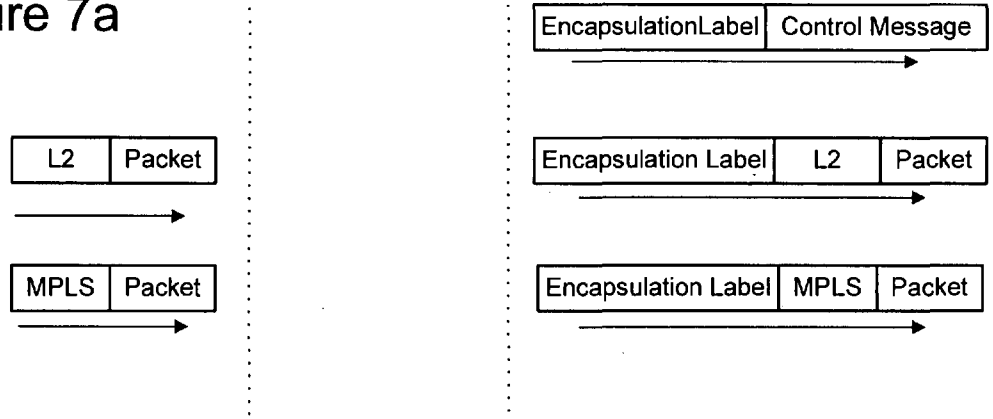


Figure 7a



60

PACKET FILTER TABLE				
PACKET FILTER (DATA INTERFACE, LABEL)	OUTGOING OPTICAL CONNECTION	ENCAPSULATION LABEL	FILTER PRIORITY	GUARANTEED QOS
PACKET FILTER-1 (PORT 1, ETHERNET VLAN 100)	SONET VCG NUMBER 3	MPLS LABEL 10000	3	X
PACKET FILTER-2 (PORT 5, ATM VCI/VPI 12/45)	SONET VCG NUMBER 3	MPLS LABEL 20000	3	Y
PACKET FILTER-3 (PORT 2, FR DLCI 900)	OPTICAL INTERFACE 1	MPLS LABEL 500	3	Z
PACKET FILTER-4 (PORT 1, MPLS LABEL 10000)	SONET VCG NUMBER 5	MPLS LABEL 10001	1	S

Figure 7b



80

CIRCUIT FILTER TABLE			
CIRCUIT FILTER (OPTICAL CONNECTION, LABEL)	OUTGOING DATA INTERFACE	OVERWRITTEN LABEL	GUARANTEED QOS
CIRCUIT FILTER-1 (SONET VCG 3, LABEL 20000)	DATA PORT 1	ETHERNET VLAN 200	X
CIRCUIT FILTER-2 (SONET VCG 3, LABEL 20001)	DATA PORT 2	ATM VCI/VPI 23/56	Y
CIRCUIT FILTER-3 (OPTICAL INTERFACE 1, LABEL 300)	DATA PORT 1	NONE	Z
CIRCUIT FILTER-4 (SONET VCG 5, LABEL 12000)	DATA PORT 23	FRAME RELAY DLCI 200	S
CIRCUIT FILTER-5 (SONET VCG 3, LABEL 3)	HOST INTERFACE	NONE	T

Figure 7c

25

SESSION TABLE			
SESSION (CONTROL MESSAGE ID)	OUTGOING OPTICAL CONNECTION	ENCAPSULATION LABEL	GUARANTEED QOS
SESSION 1 (TCP SRC PORT 1345)	SONET VCG NUMBER 3	MPLS LABEL 3	X
SESSION 2 (TCP SRC PORT 3456)	SONET VCG NUMBER 3	MPLS LABEL 3	Y
SESSION 3 (UDP PORT 1998)	OPTICAL INTERFACE 1	MPLS LABEL 10000	Z

Figure 7d

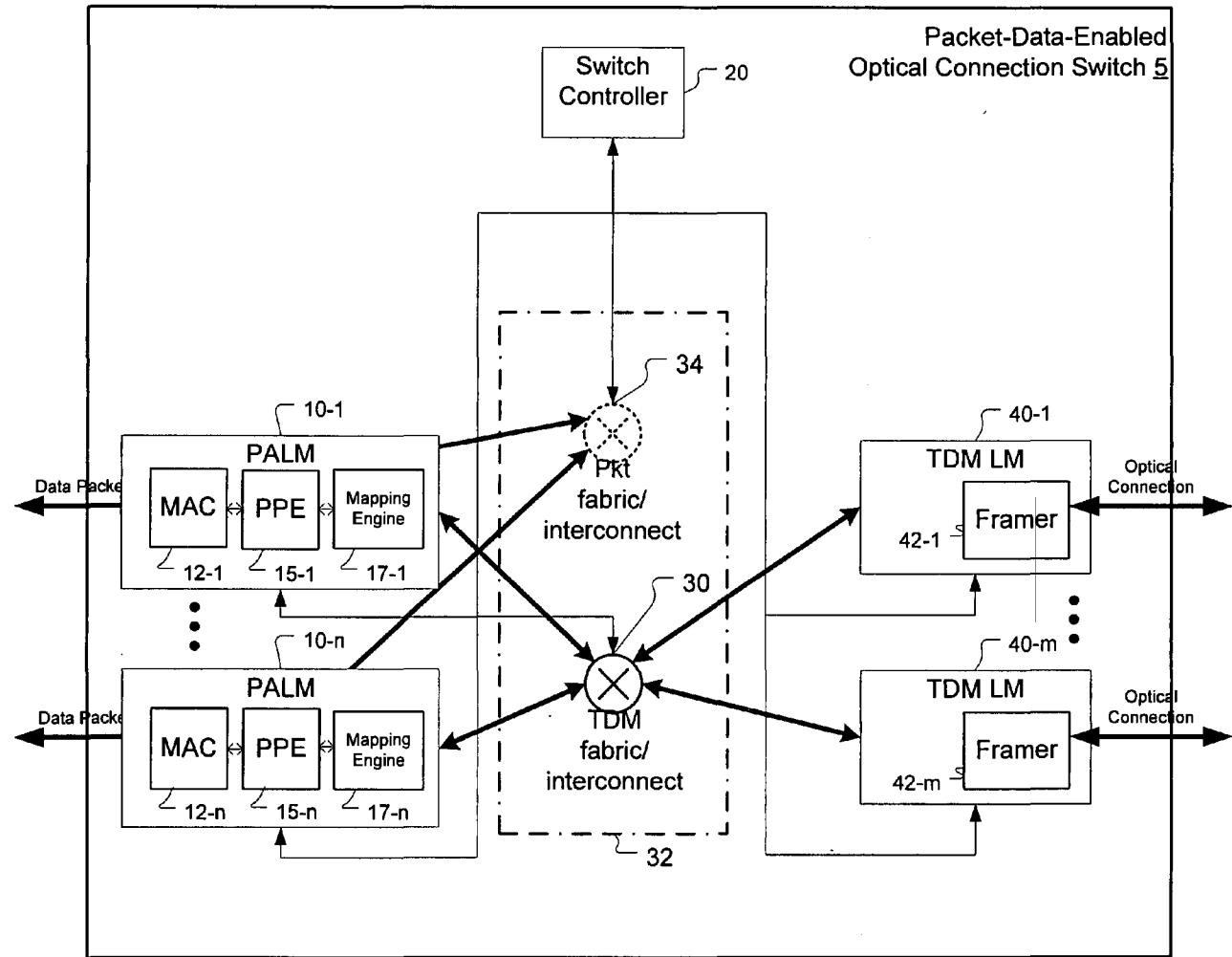


Figure 8

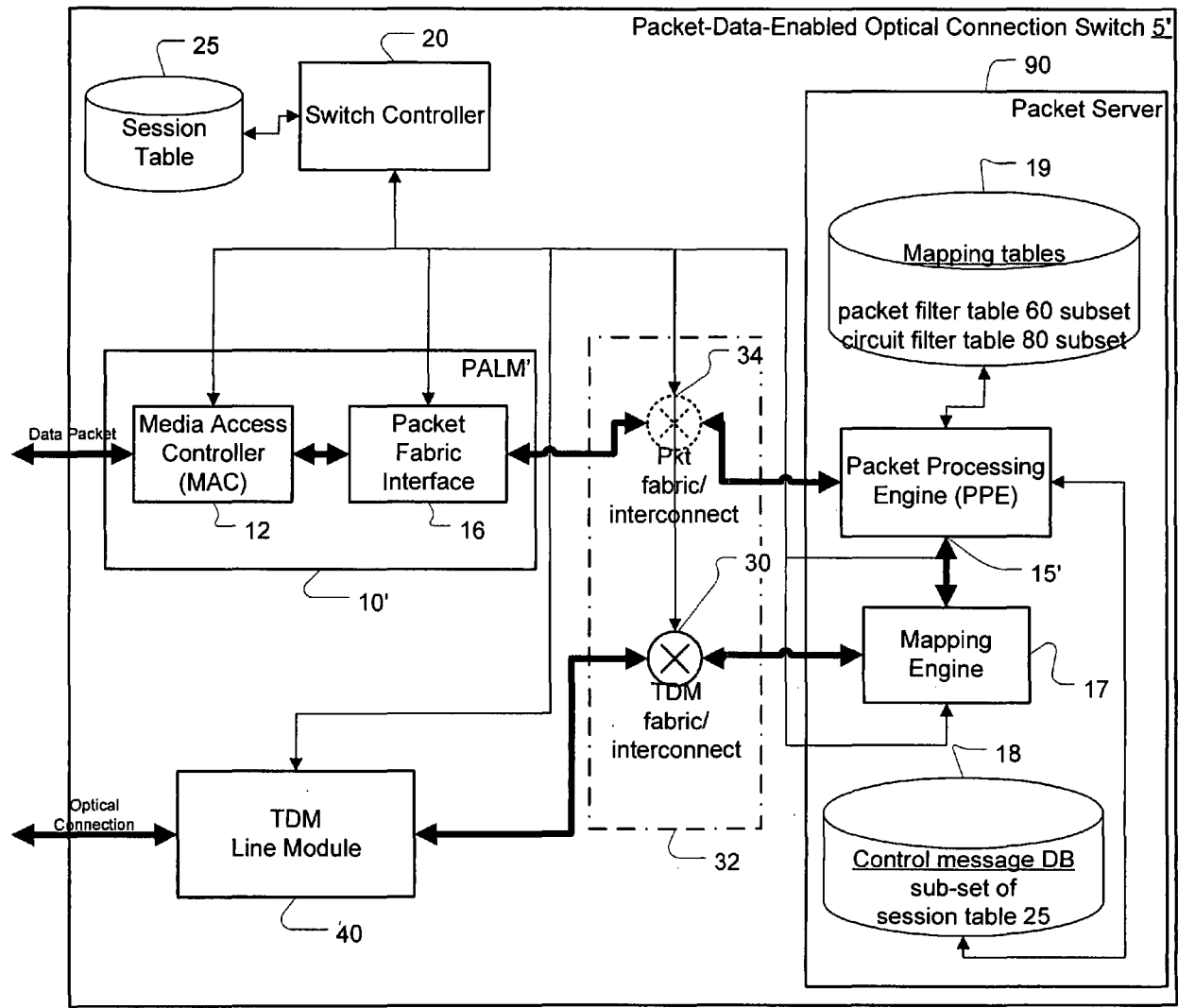


Figure 9

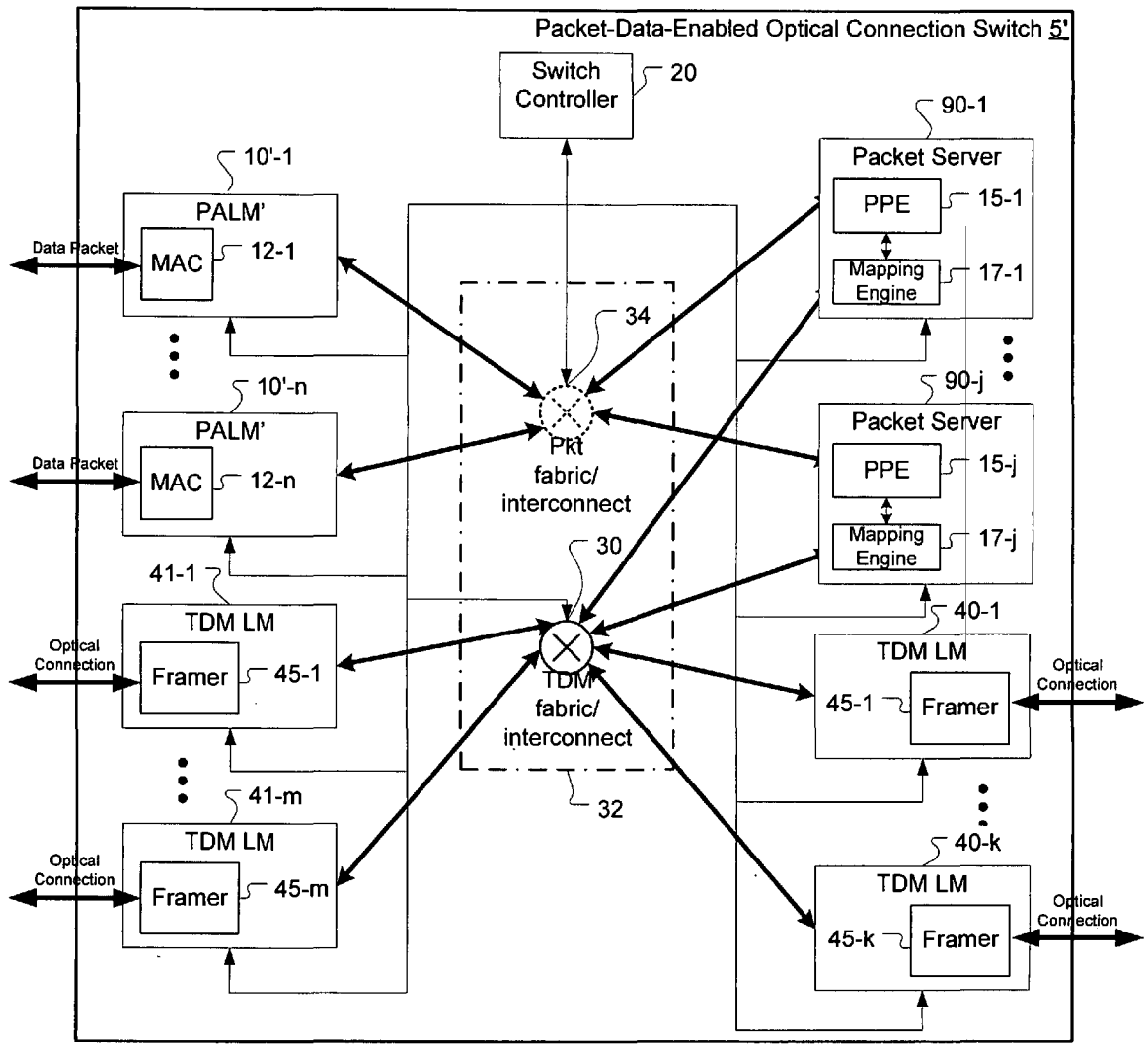


Figure 10

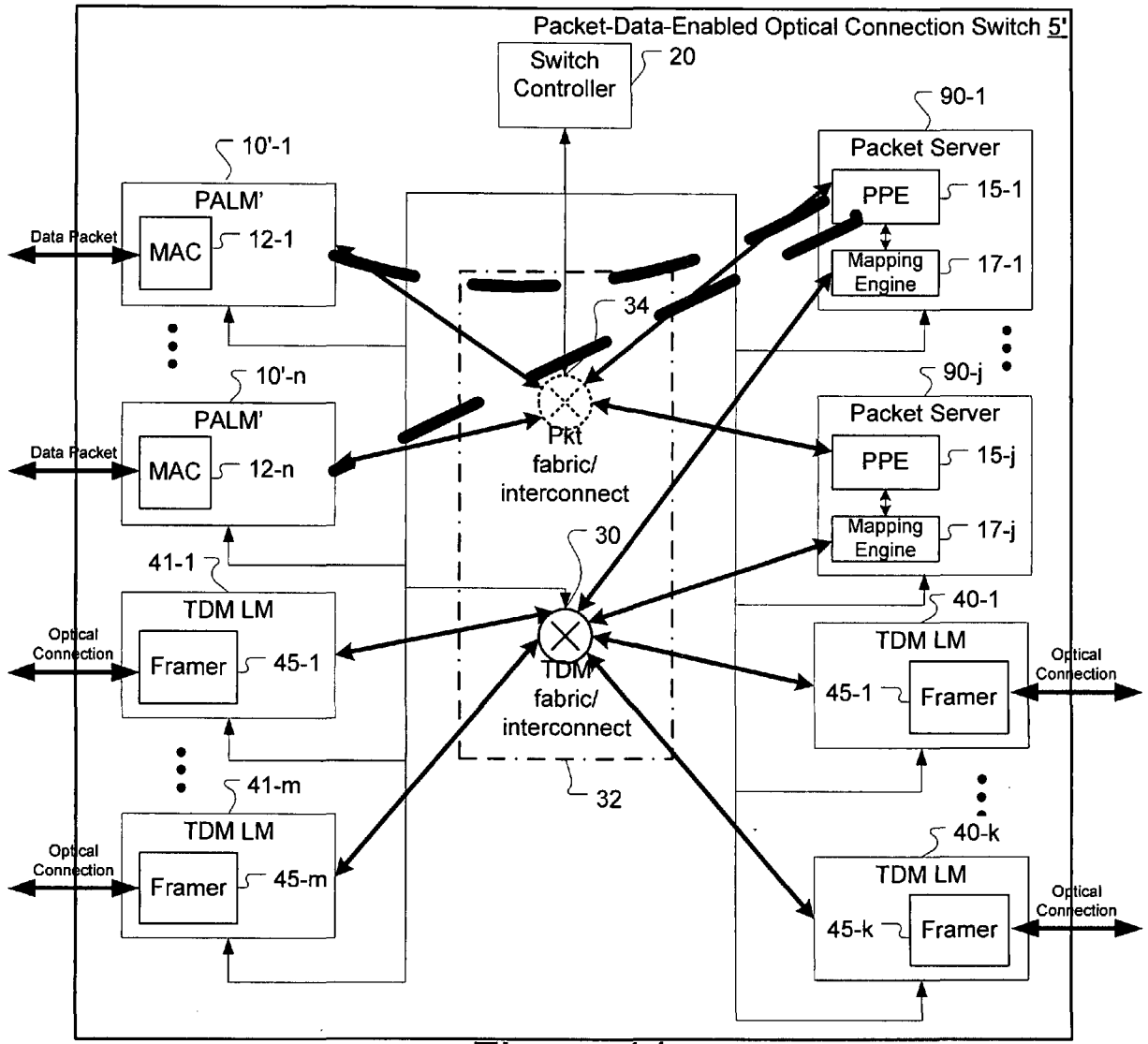


Figure 11

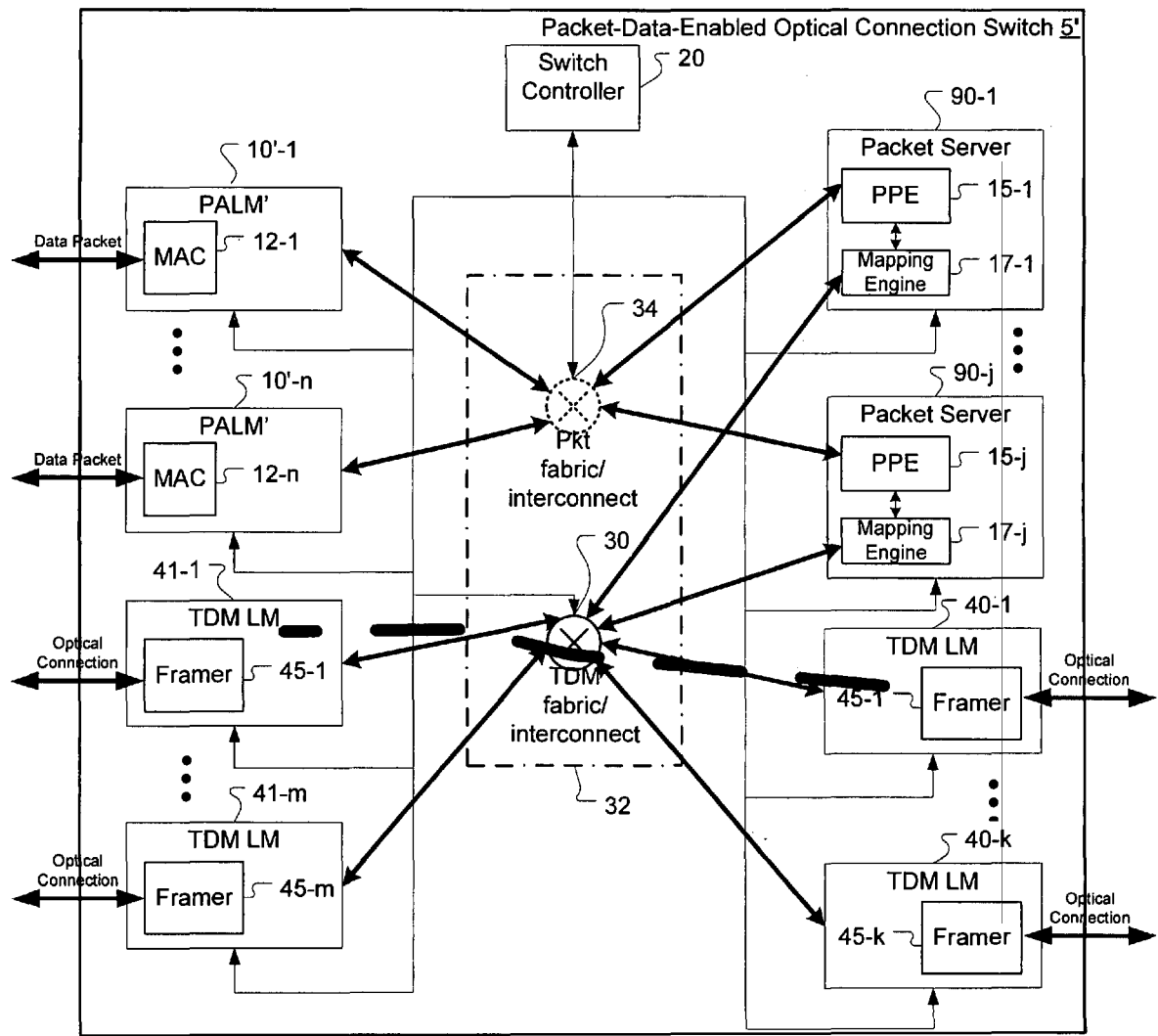


Figure 12

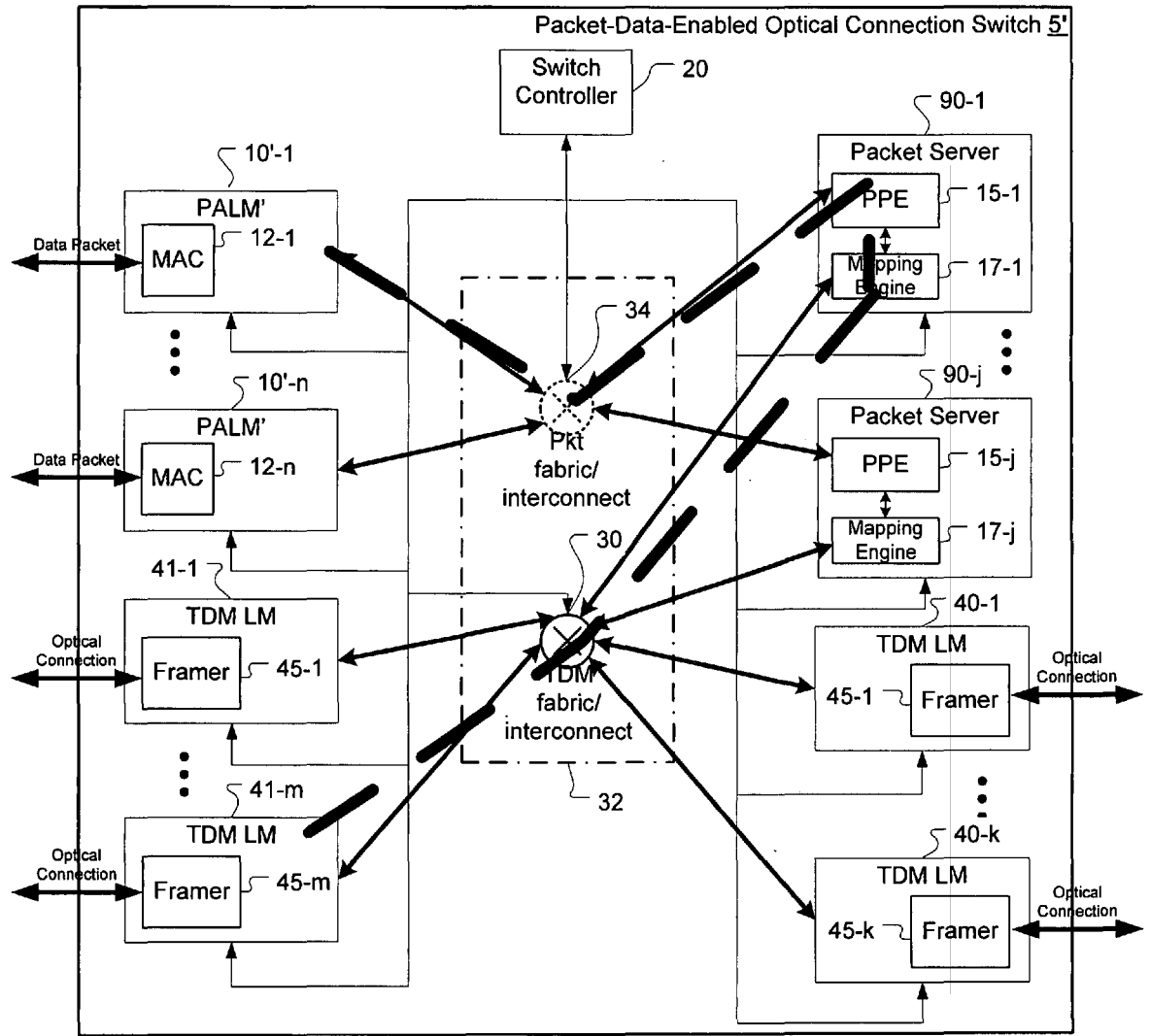


Figure 13



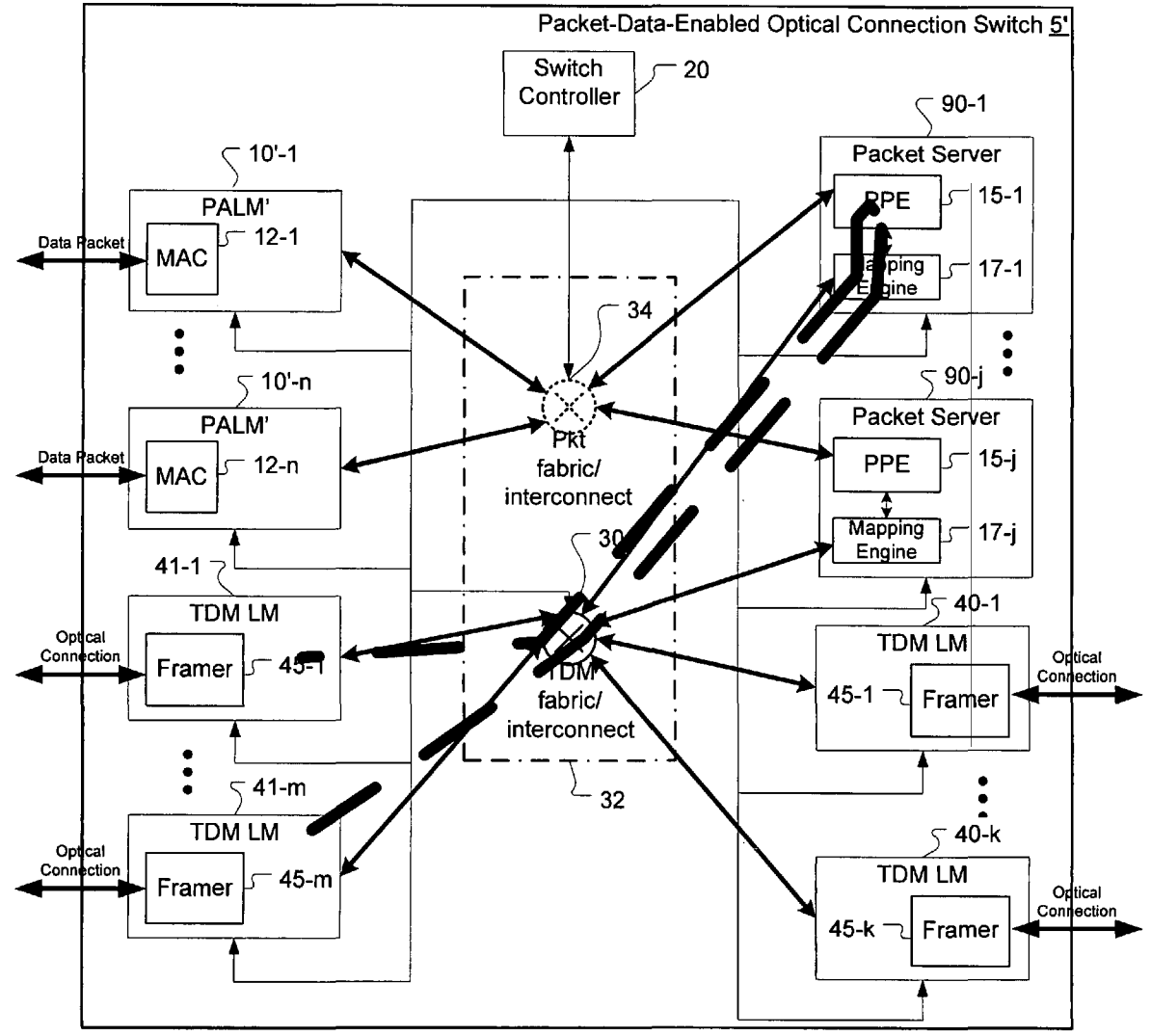


Figure 13a

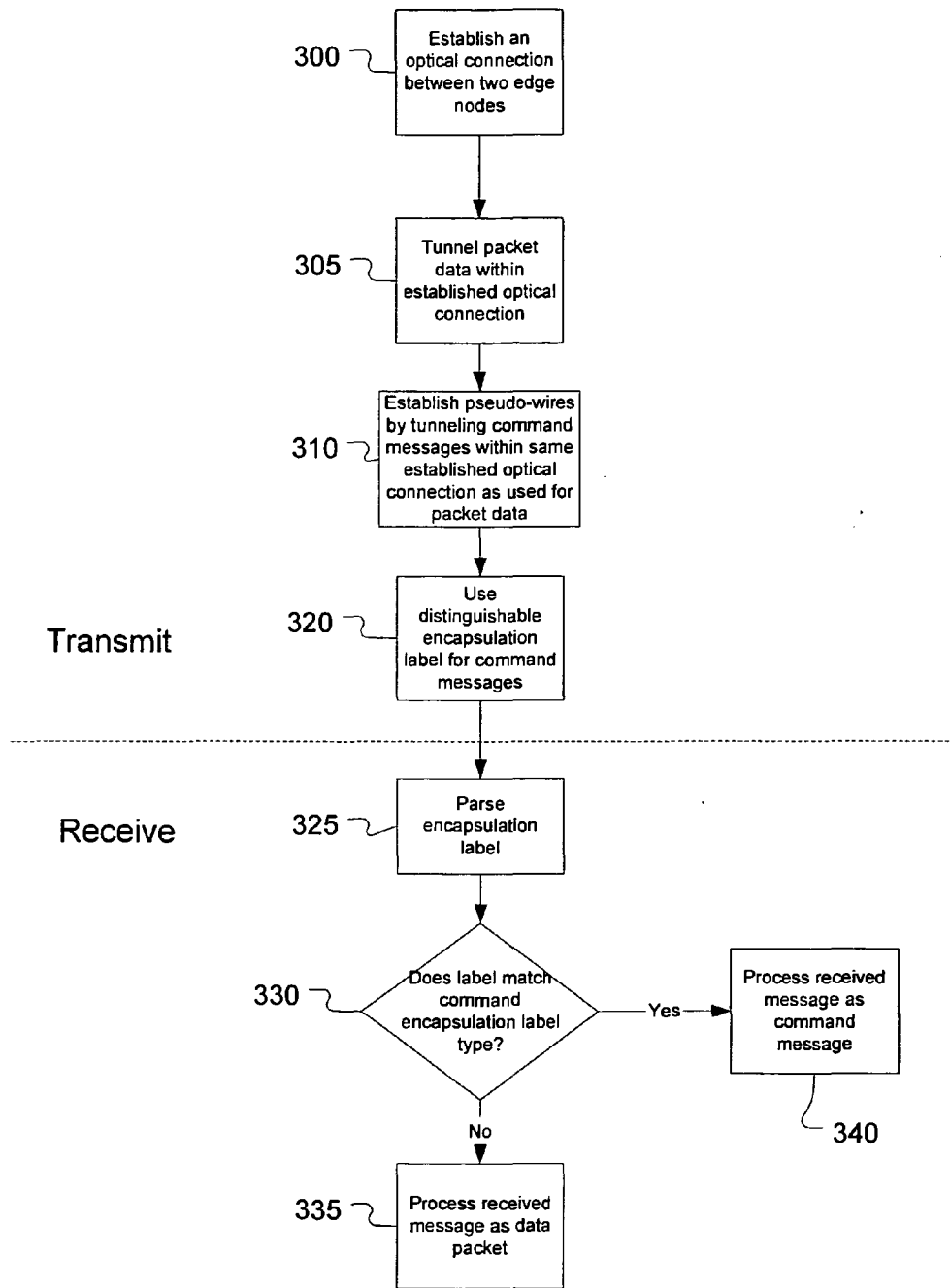


Figure 14

Figure 15a

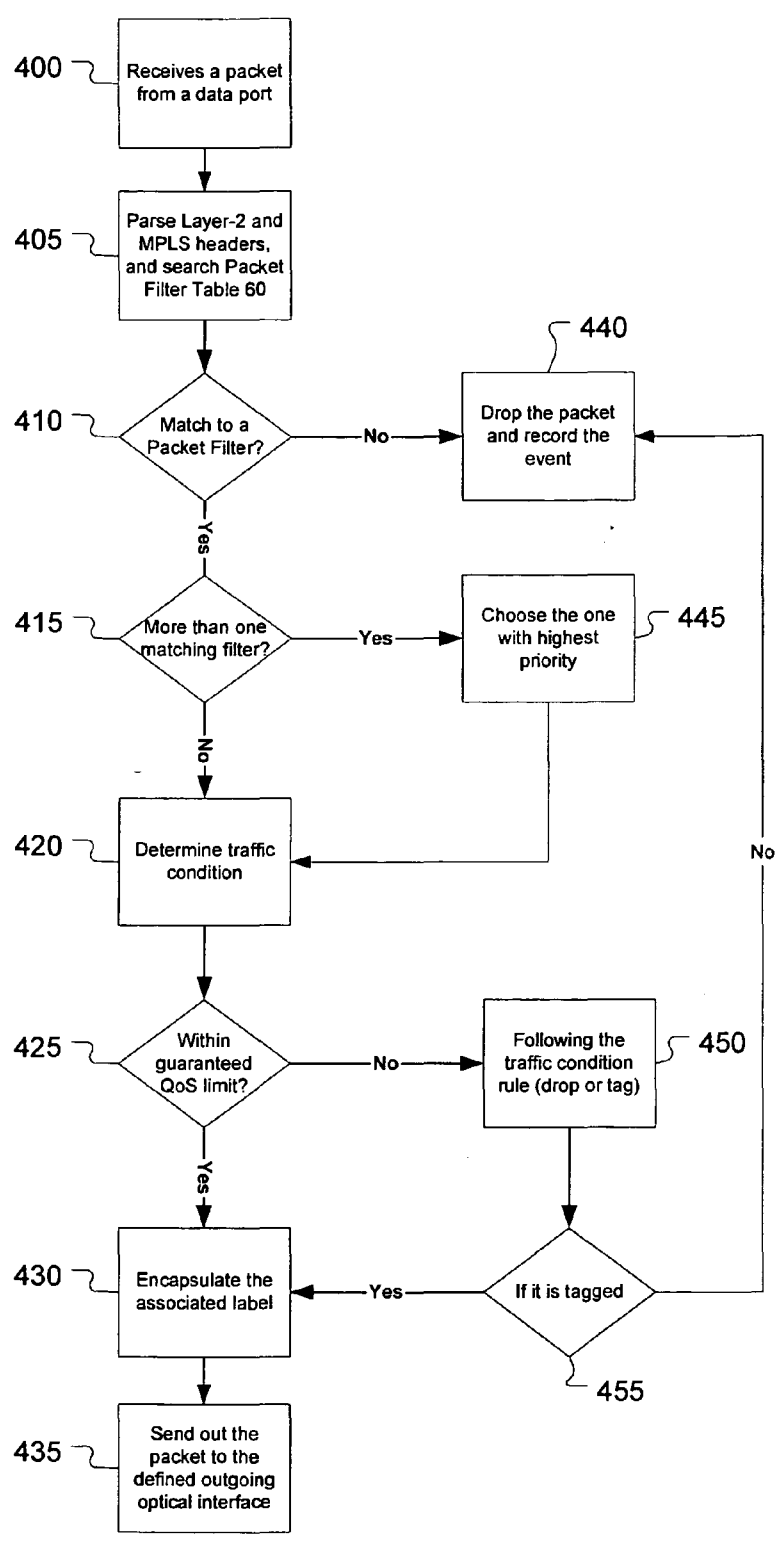


Figure 15b

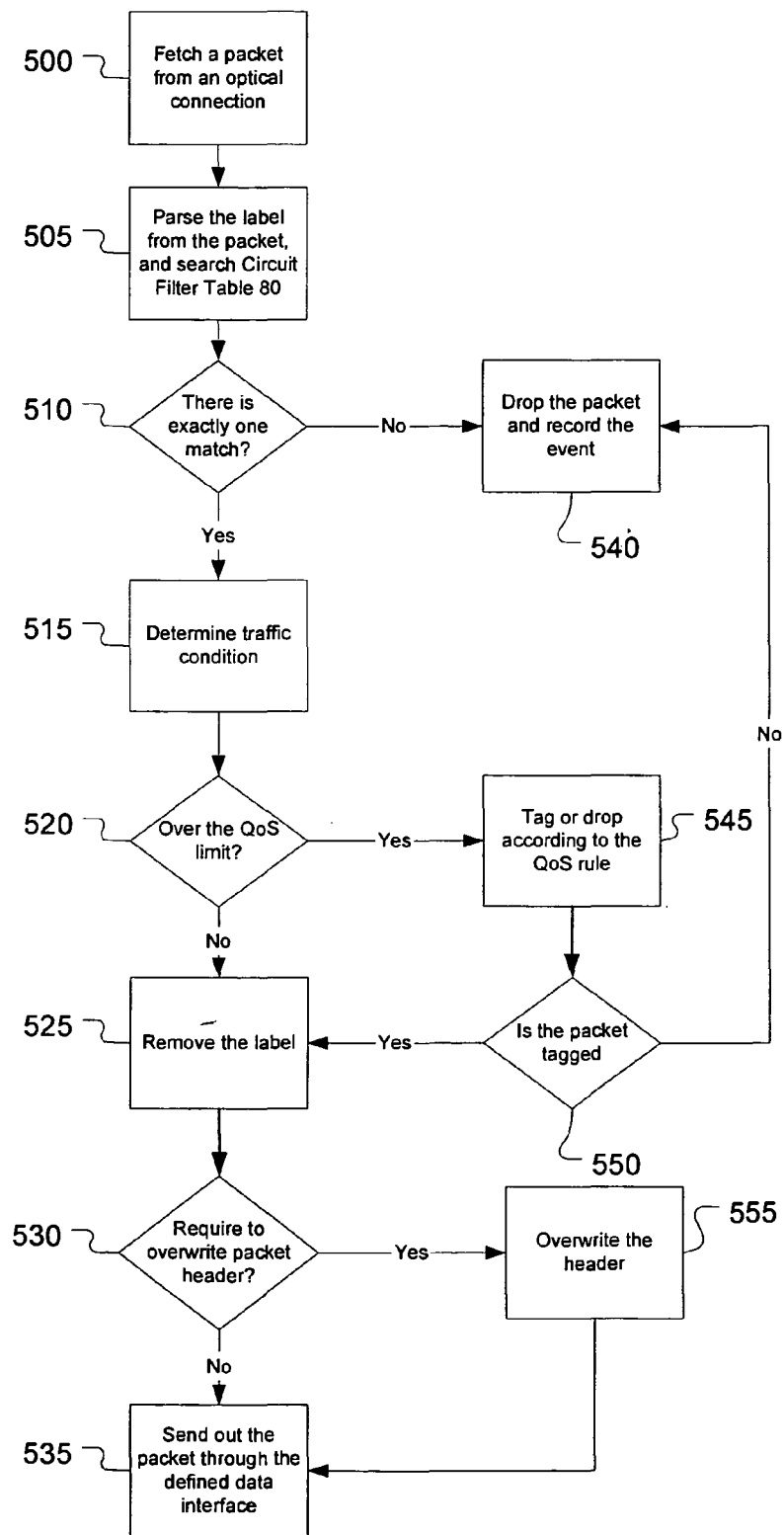
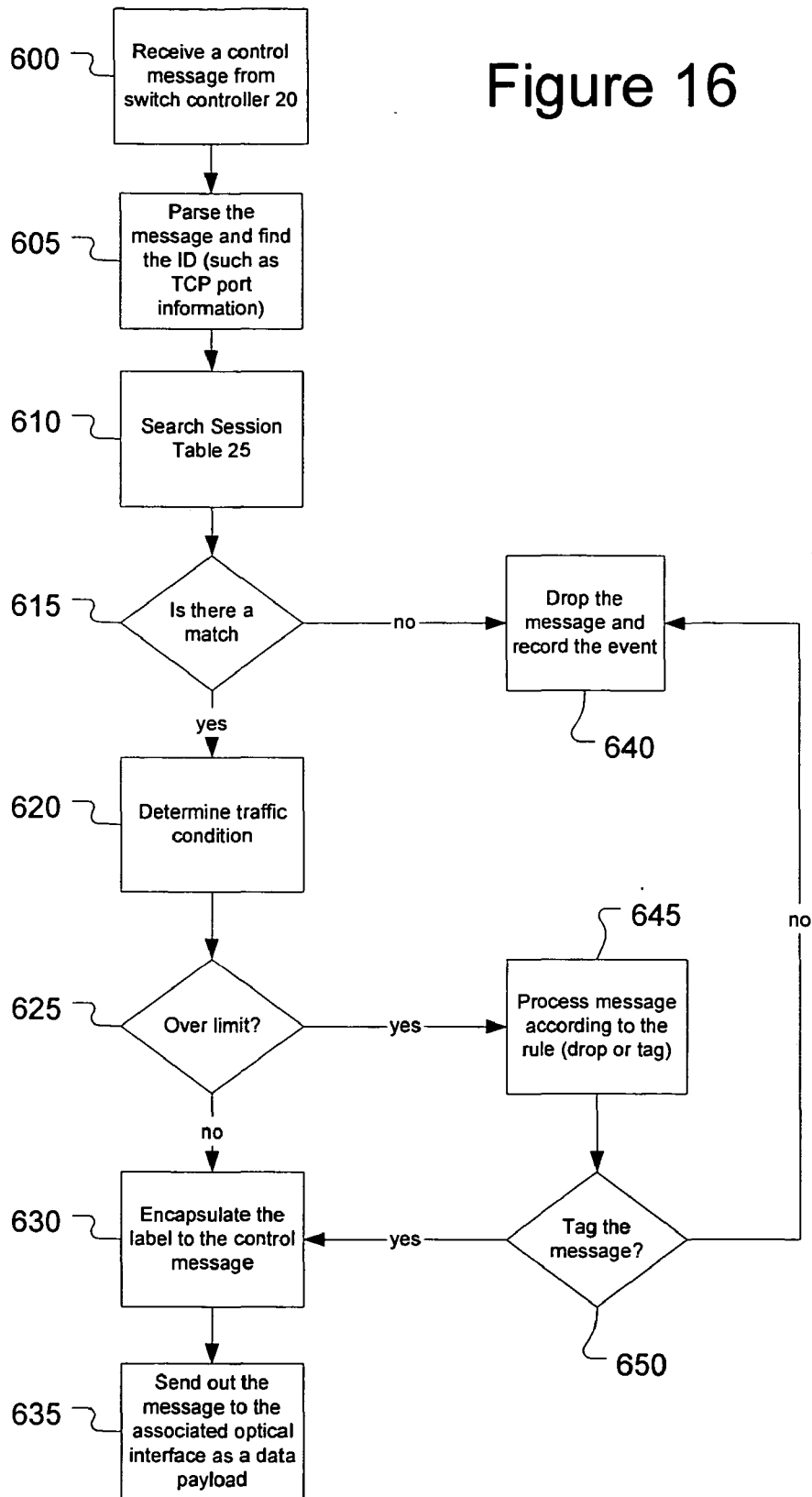


Figure 16



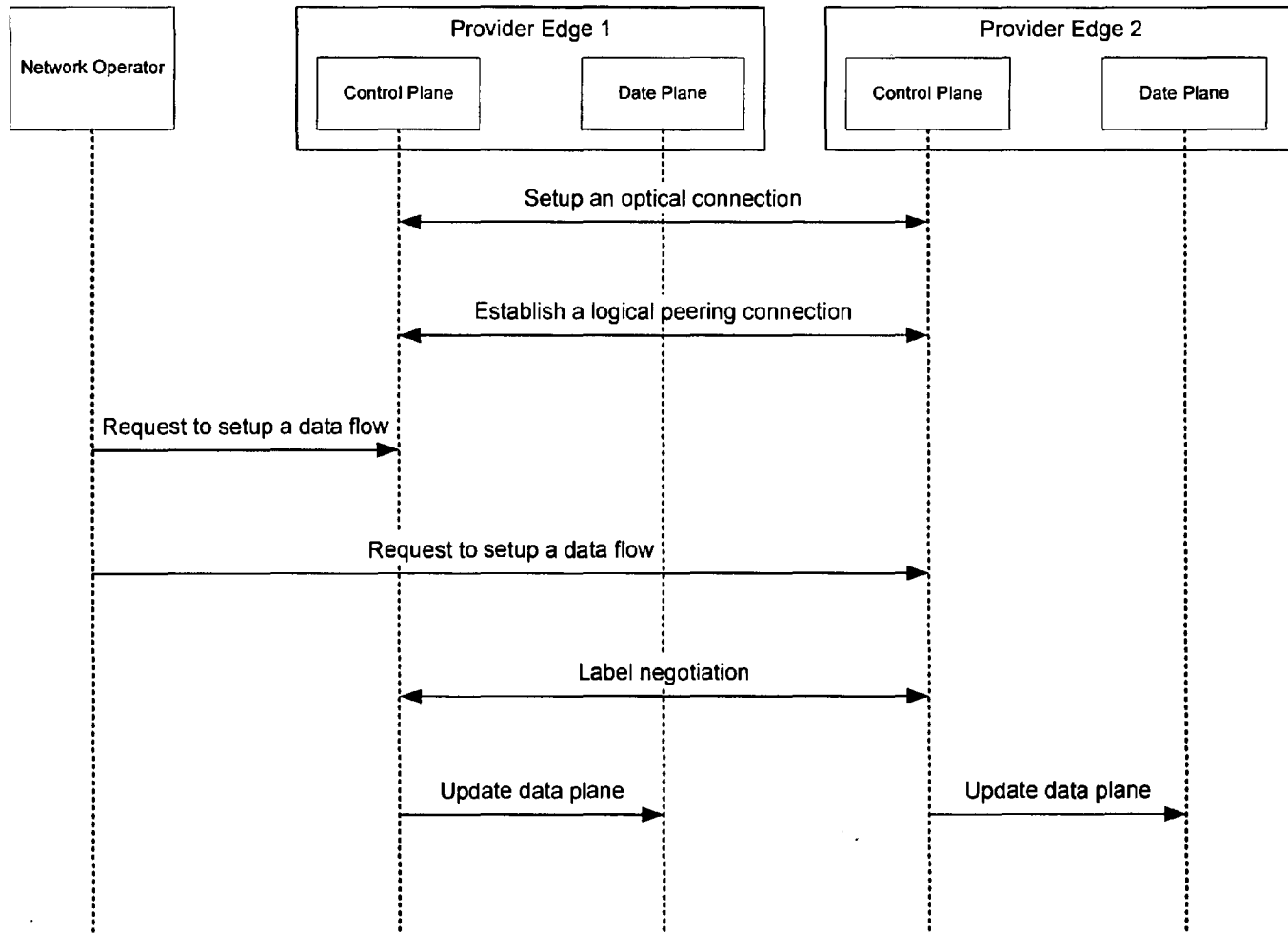


Figure 17

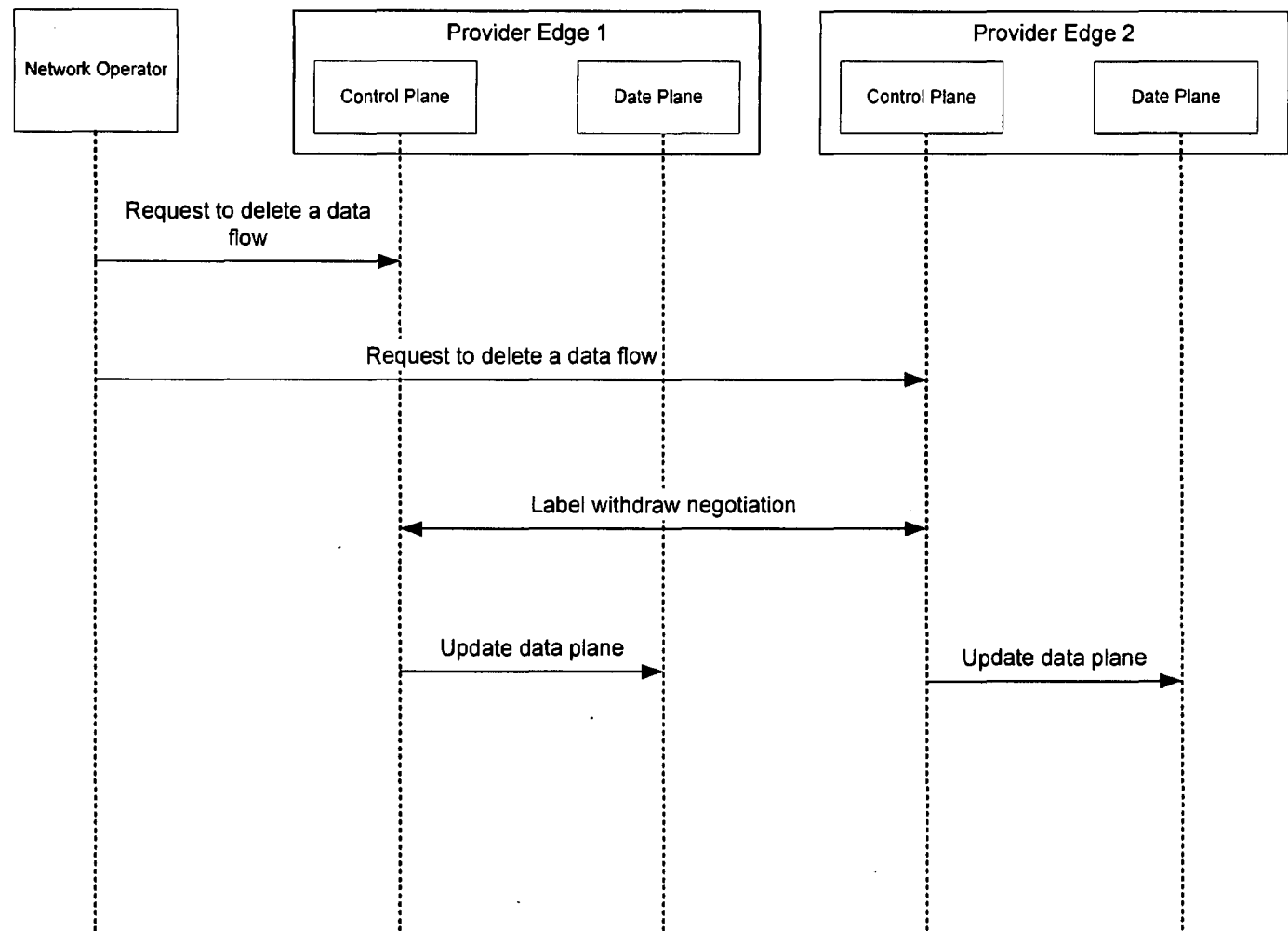


Figure 18

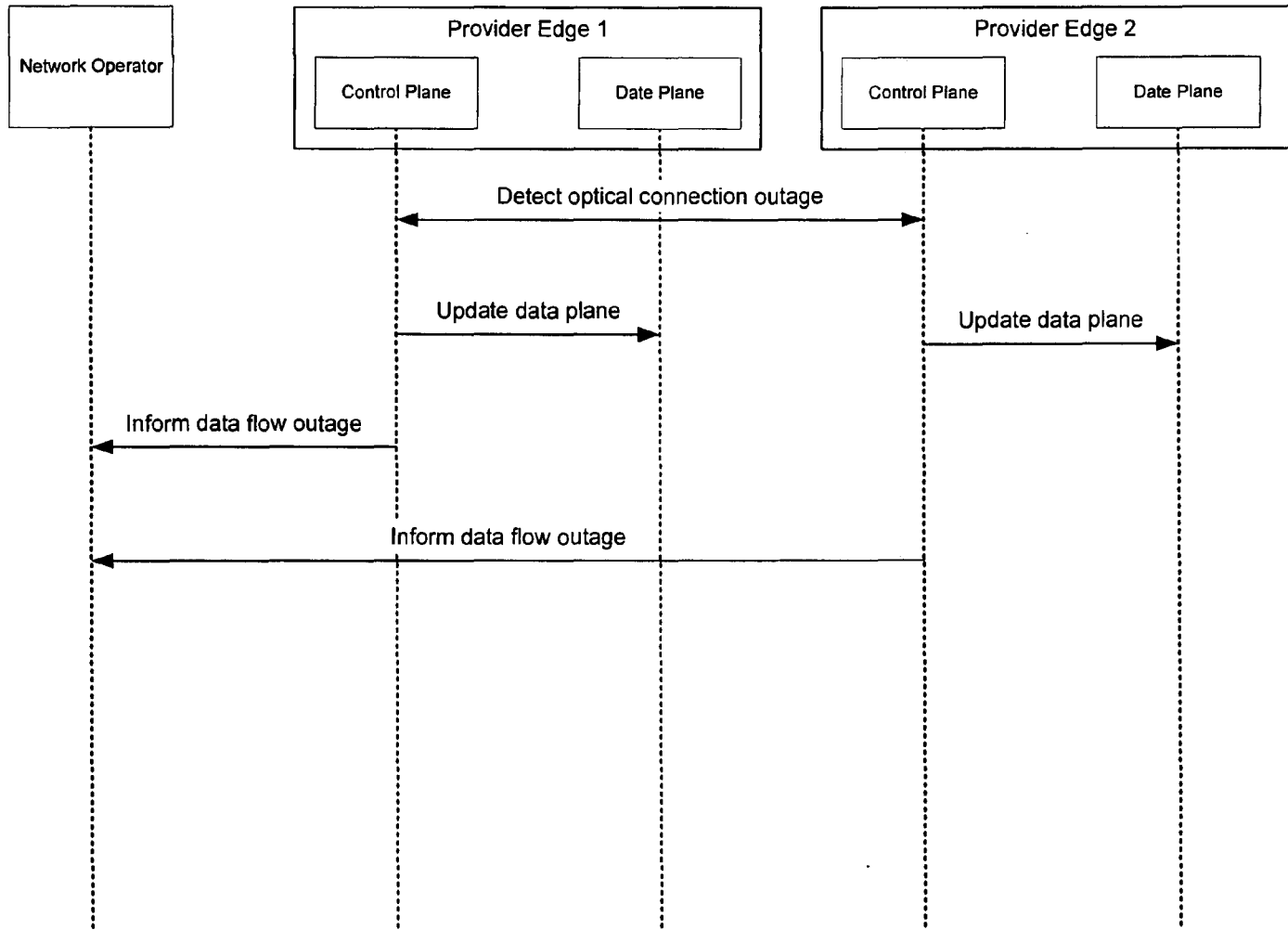


Figure 19



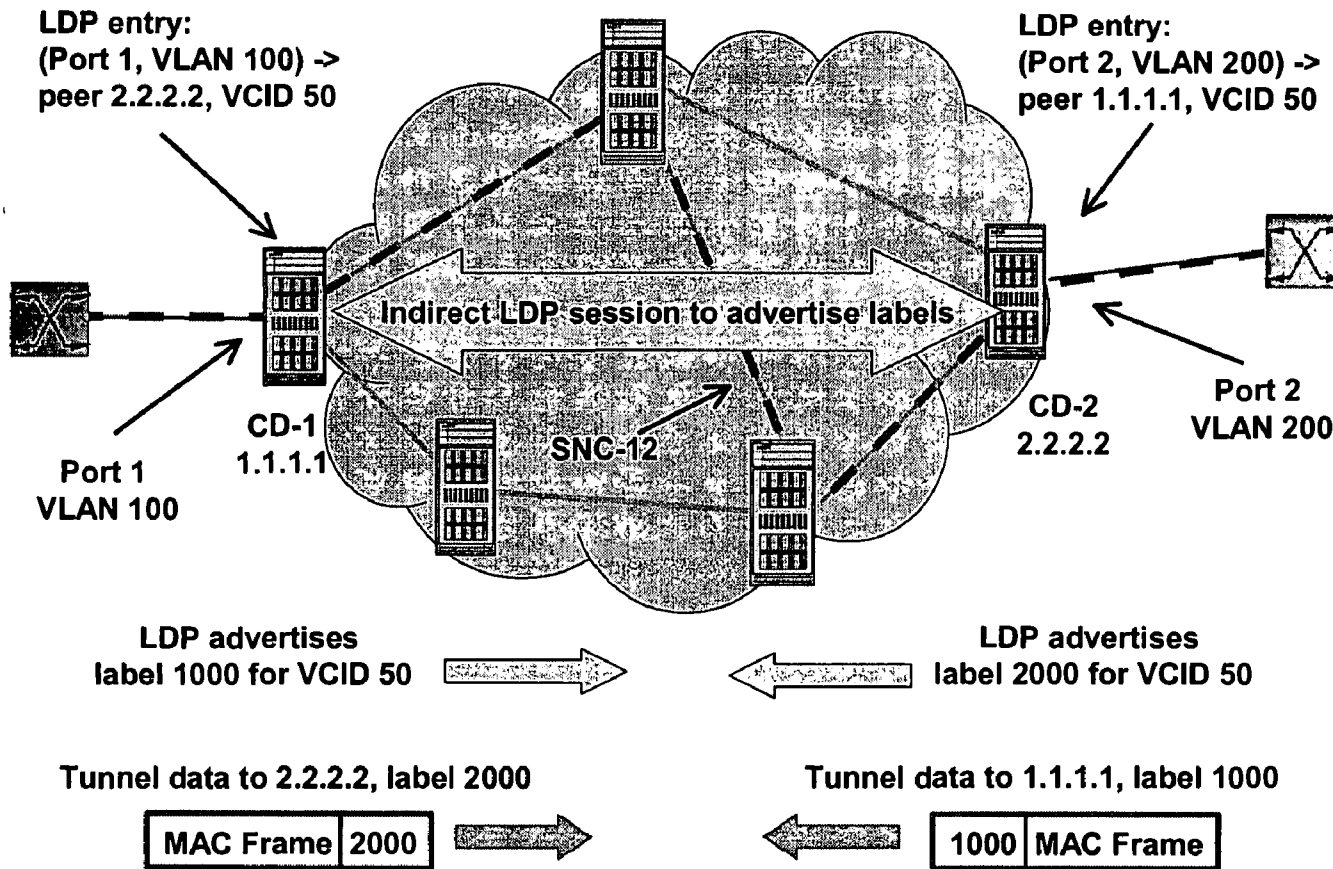


Figure 20

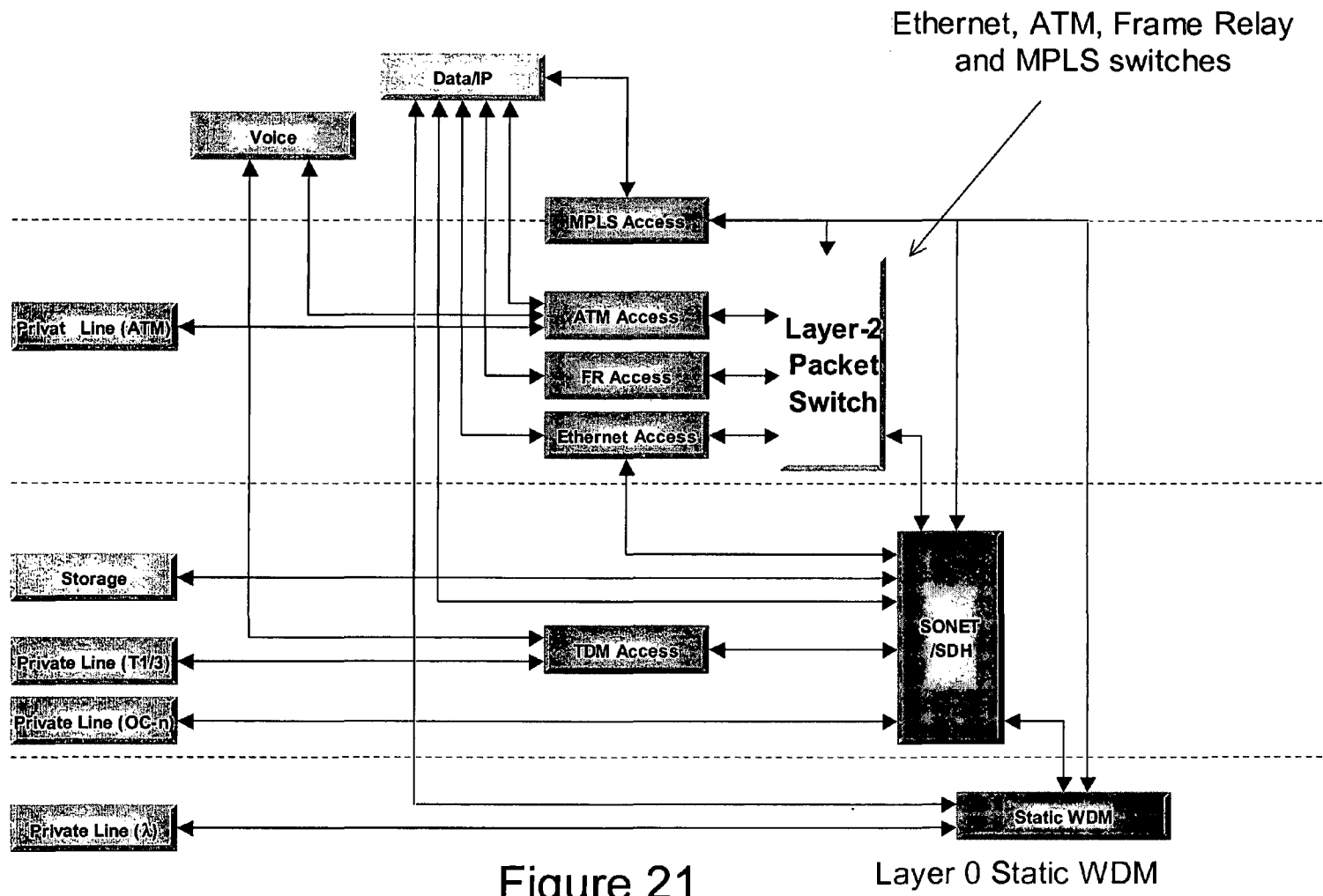


Figure 21  
(Conventional)

Ethernet, ATM, Frame Relay  
and MPLS switches

Layer 0 Static WDM

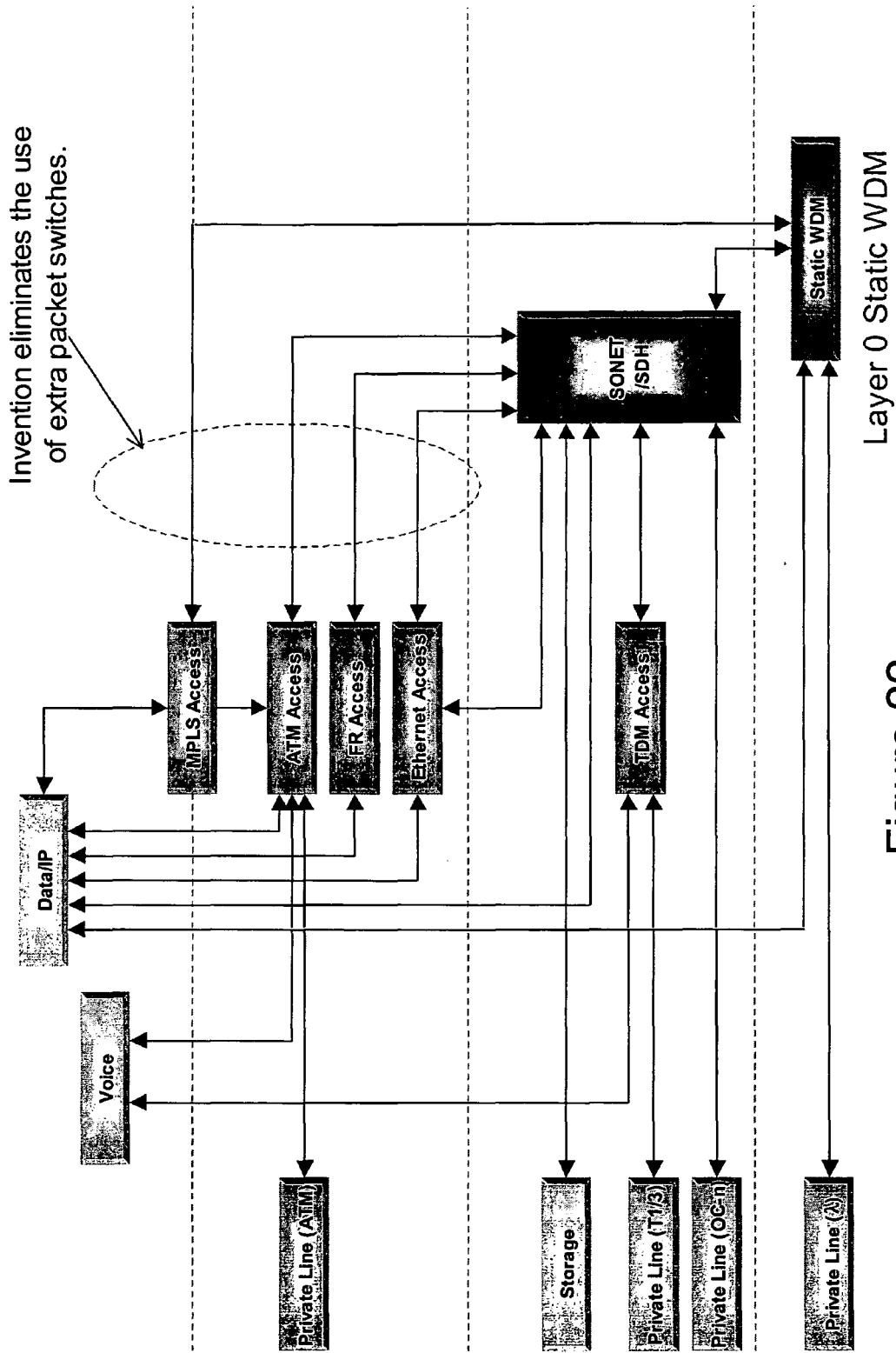


Figure 22

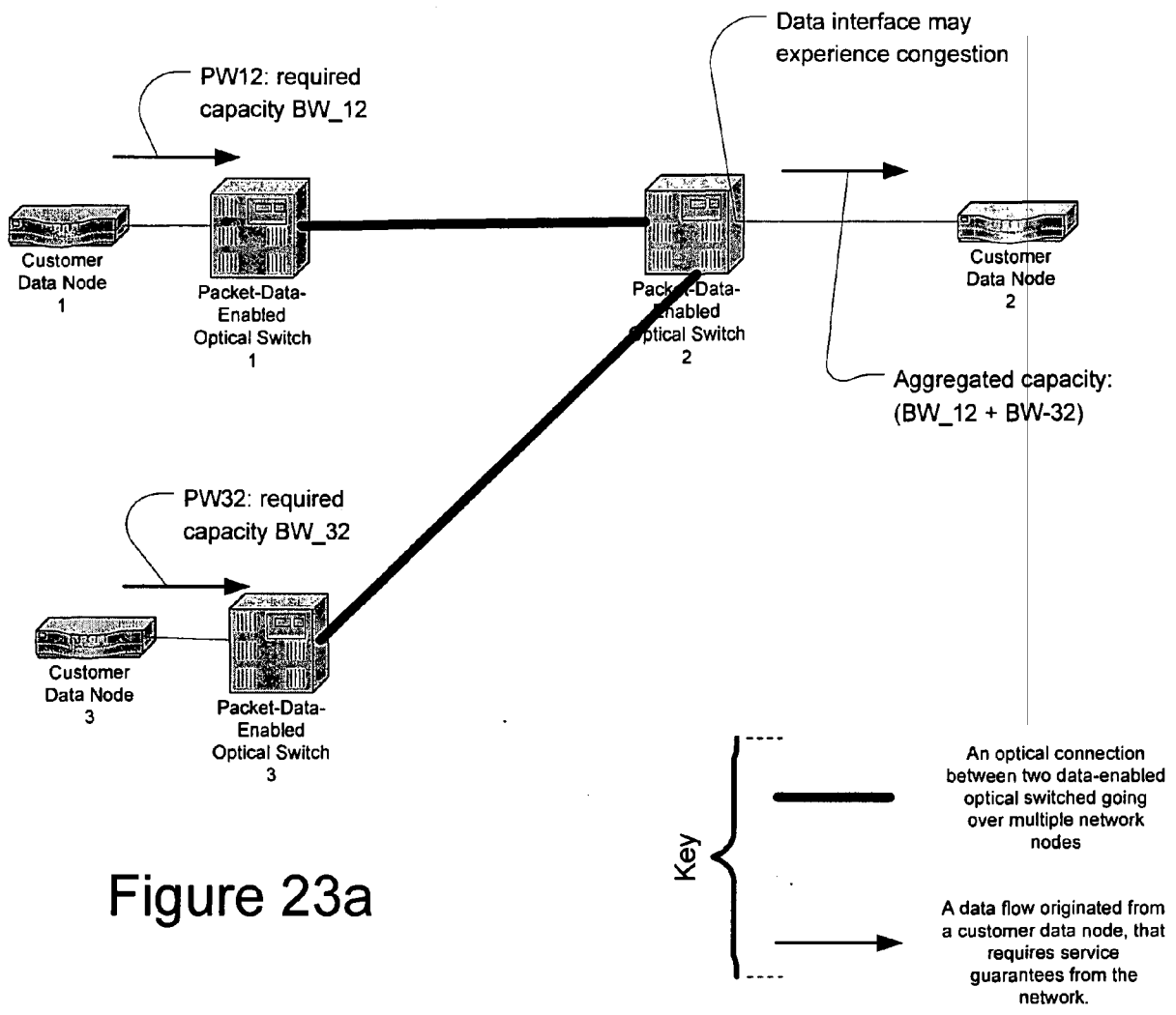


Figure 23a

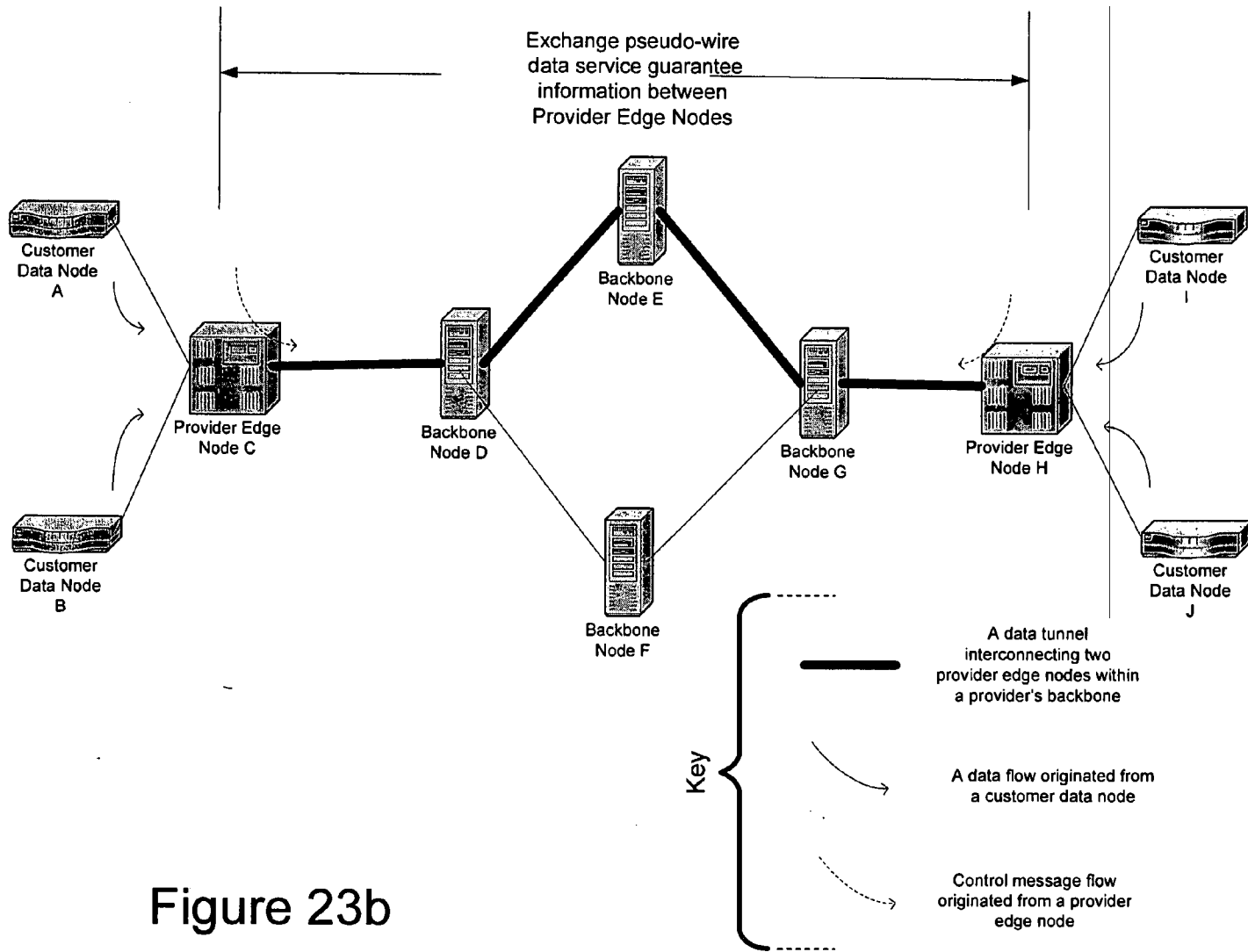


Figure 23b

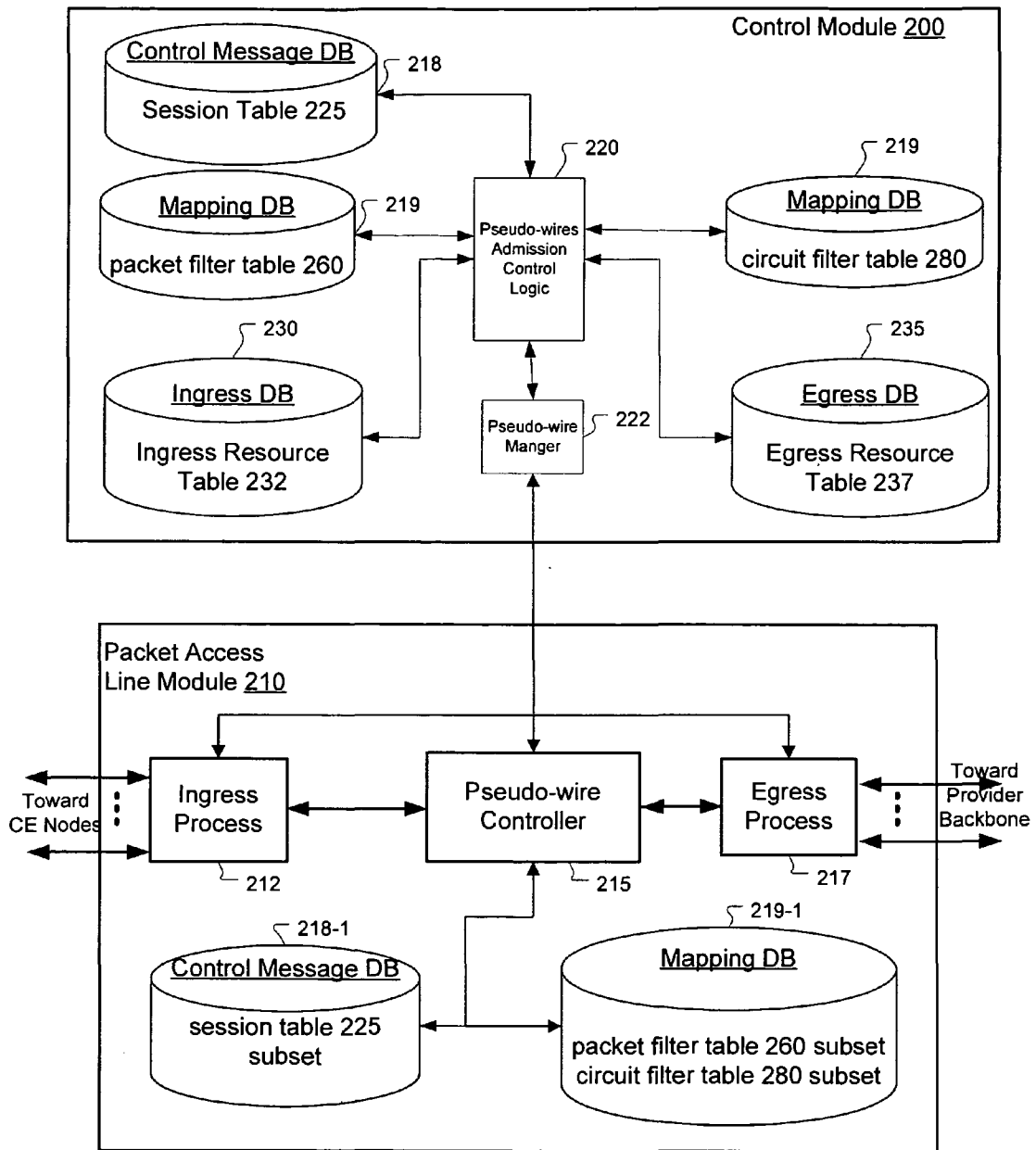


Figure 24

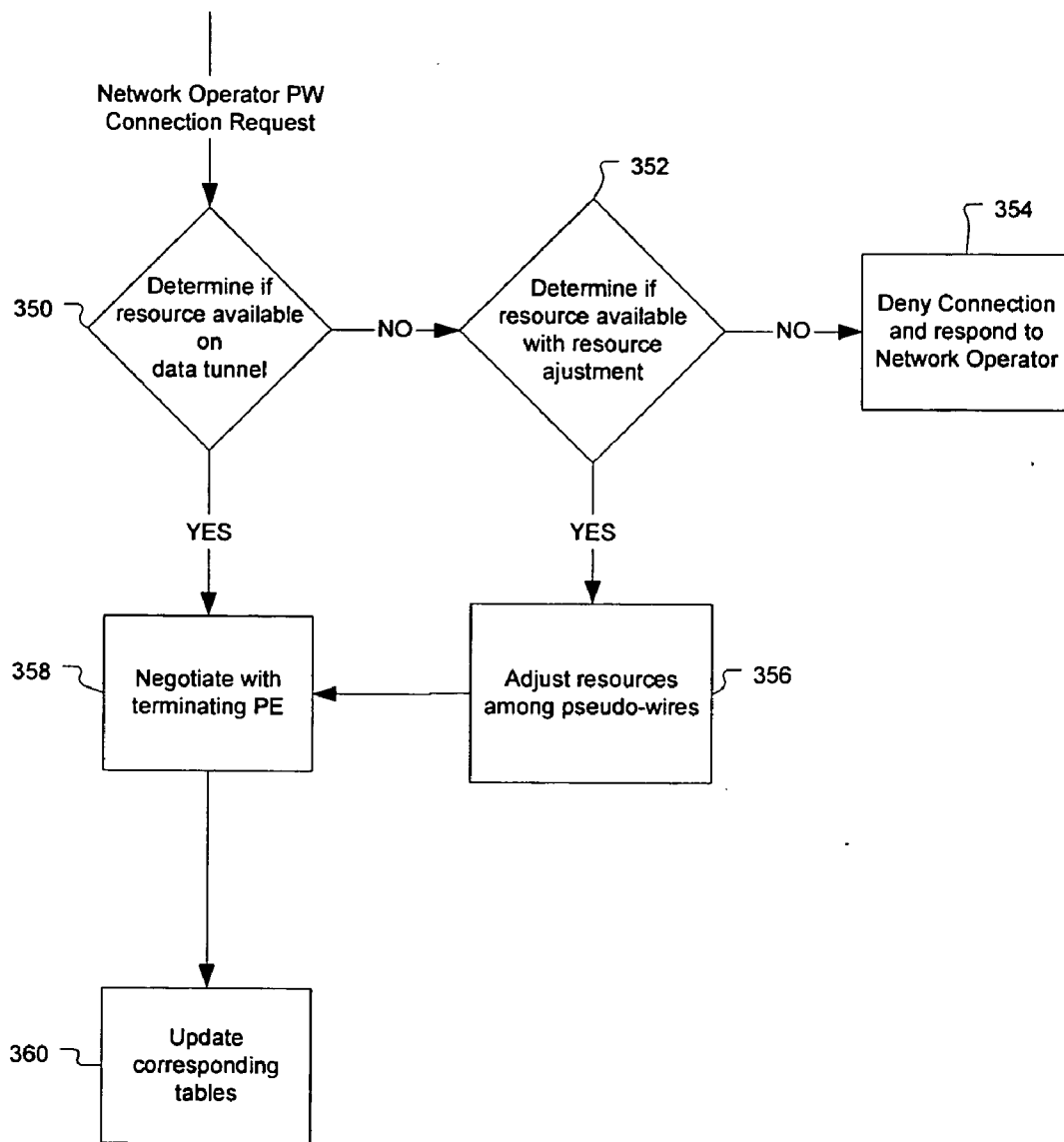


Figure 25a

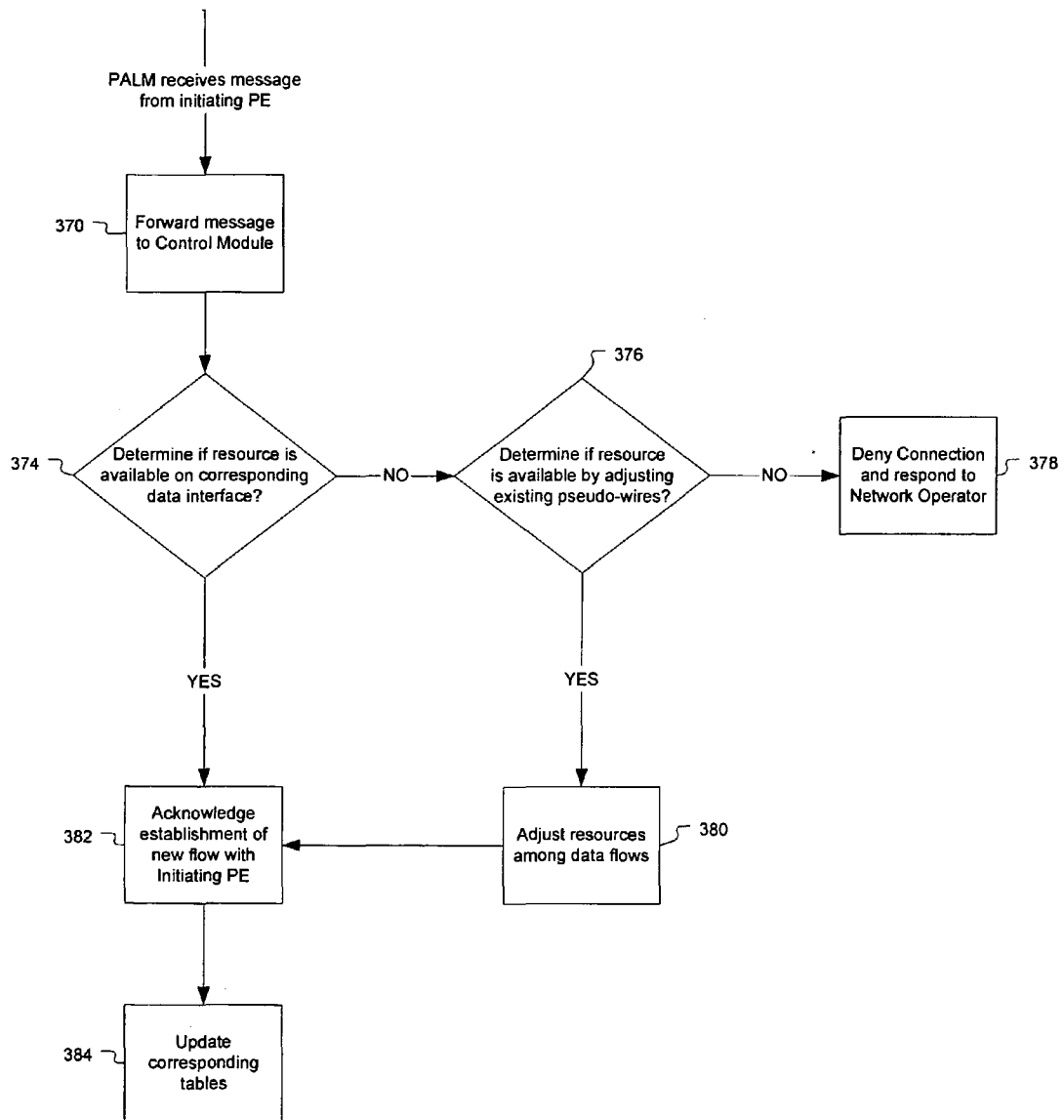


Figure 25b



260

PACKET FILTER TABLE						
PACKET FILTER (DATA INTERFACE, LABEL)	DATA TUNNEL	ENCAPSULATION LABEL	CIR	CLASS	SETUP PRIORITY	HOLDING PRIORITY
PACKET FILTER-1 (PORT 1, ETHERNET VLAN 100)	SONET VCG NUMBER 3	MPLS LABEL 10000	50 Mb/Sec	AF-1	3	5
PACKET FILTER-2 (PORT 5, ATM VCI/VPI 12/45)	ROUTER INTERFACE 5	MPLS LABEL 20000	8 Mb/Sec	EF	3	5
PACKET FILTER-3 (PORT 2, FR DLCI 900)	SONET VCG NUMBER 3	MPLS LABEL 500	10 Mb/Sec	AF-1	3	5
PACKET FILTER-4 (PORT 10)	ETHERNET INTERFACE 12	MPLS LABEL 50001	1 Gb/Sec	AF-3	5	5

Figure 26

280

CIRCUIT FILTER TABLE					
CIRCUIT FILTER (DATA TUNNEL, LABEL)	OUTGOING DATA INTERFACE	CIR	CLASS	SETUP PRIORITY	HOLDING PRIORITY
CIRCUIT FILTER-1 (SONET VCG 3, LABEL 20000)	DATA PORT 1	50 Mb/Sec	AF-1	3	3
CIRCUIT FILTER-2 (SONET VCG 3, LABEL 20001)	DATA PORT 2	8 Mb/Sec	EF	3	3
CIRCUIT FILTER-3 (OPTICAL INTERFACE 1, LABEL 300)	DATA PORT 10	10 Gb/Sec	AF-3	3	3
CIRCUIT FILTER-4 (SONET VCG 5, LABEL 12000)	DATA PORT 1	100 Mb/Sec	AF-1	1	1
CIRCUIT FILTER-5 (SONET VCG 3, LABEL 3)	DATA PORT 2	1 Gb/Sec	AF-1	5	5

Figure 27

225

SESSION TABLE				
SESSION (CONTROL MESSAGE ID)	OUTGOING DATA TUNNEL	ENCAPSULATION LABEL	CIR	CLASS
SESSION 1 (TCP SRC PORT 1345)	SONET VCG NUMBER 3	MPLS LABEL 3	1 Mb/Sec	EF
SESSION 2 (TCP SRC PORT 3456)	MPLS LSP 8	MPLS LABEL 3	2 Mb/Sec	EF
SESSION 3 (TCP SRC PORT 1998)	OPTICAL INTERFACE 1	MPLS LABEL 10000	N/A	EF

Figure 28

232

INGRESS RESOURCE TABLE					
DATA TUNNEL	PHYSICAL BANDWIDTH	AVAILABLE TOTAL BANDWIDTH	AVAILABLE BANDWIDTH CLASS 1	AVAILABLE BANDWIDTH CLASS ...	AVAILABLE BANDWIDTH CLASS N
MPLS LSP with label 45	80 Mb/Sec	70 Mb/Sec	20 Mb/sec	...	30 Mb/sec
SONET VCG 4	100 Mb/Sec	50 Mb/Sec	50 Mb/Sec	...	0
DEDICATED POS CONNECTION	45 Mb/Sec	12 Mb/Sec	1.5 Mb/src	...	3.0 Mb/src

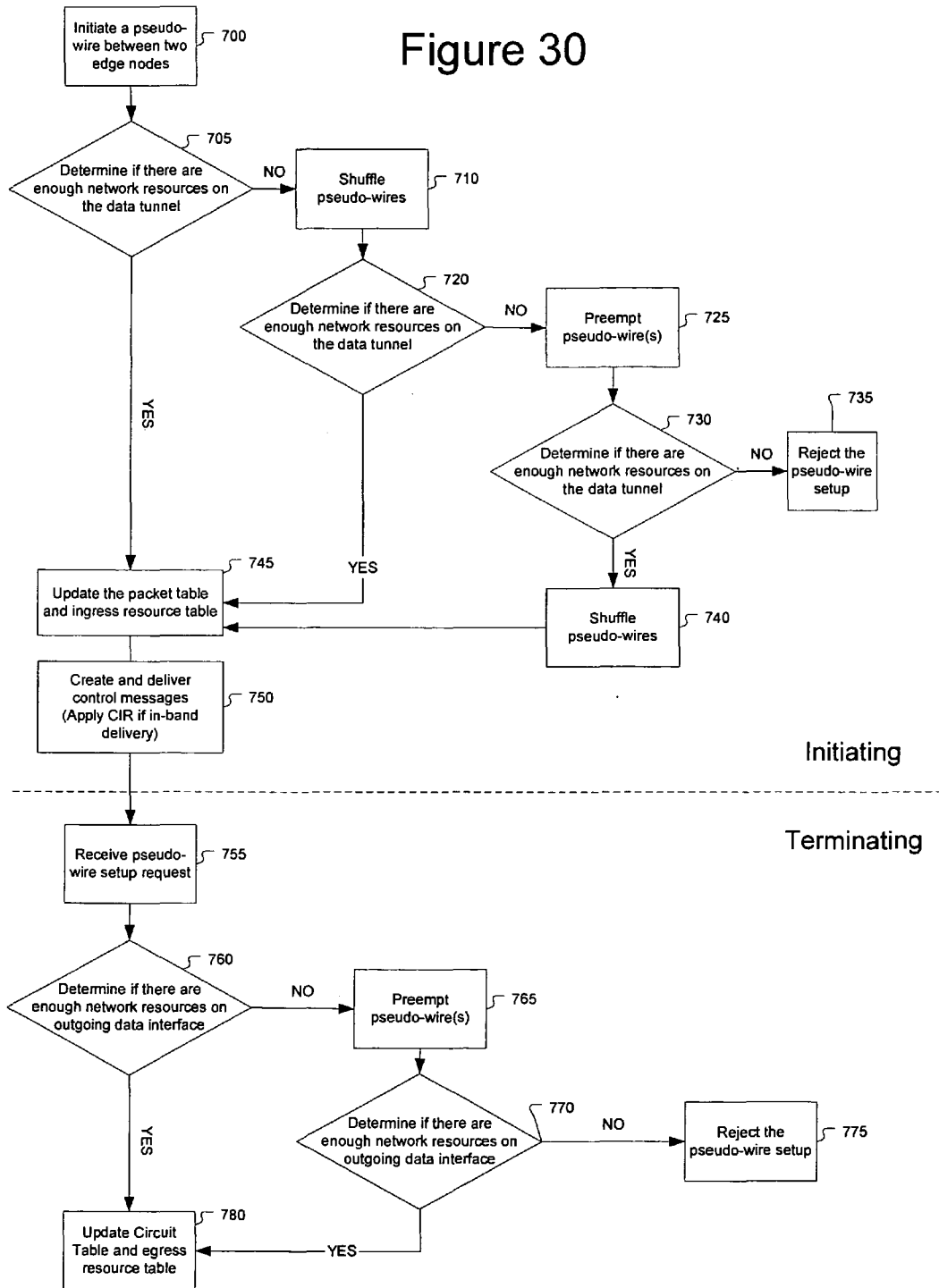
Figure 29a

237

EGRESS RESOURCE TABLE					
DATA INTERFACE	PHYSICAL BANDWIDTH	AVAILABLE TOTAL BANDWIDTH	AVAILABLE BANDWIDTH CLASS 1	AVAILABLE BANDWIDTH CLASS ...	AVAILABLE BANDWIDTH CLASS N
ATM Interface	150 Mb/Sec	70 Mb/Sec	20 Mb/sec	...	30 Mb/sec
Ethernet Interface	100 Mb/Sec	50 Mb/Sec	50 Mb/Sec	...	0
DS3 Interface	45 Mb/Sec	12 Mb/Sec	1.5 Mb/src	...	3.0 Mb/src

Figure 29b

Figure 30



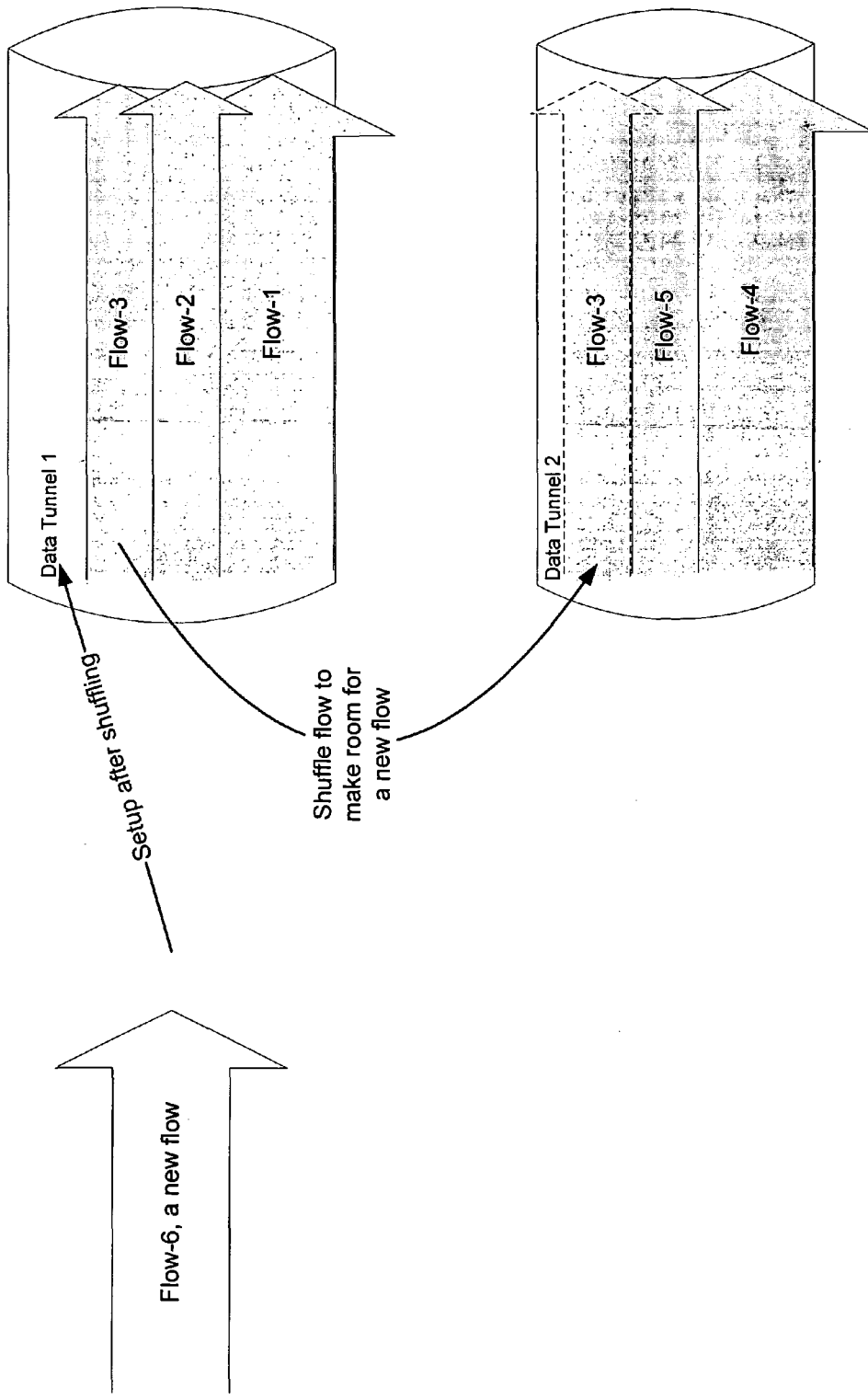


Figure 31

# Figure 32

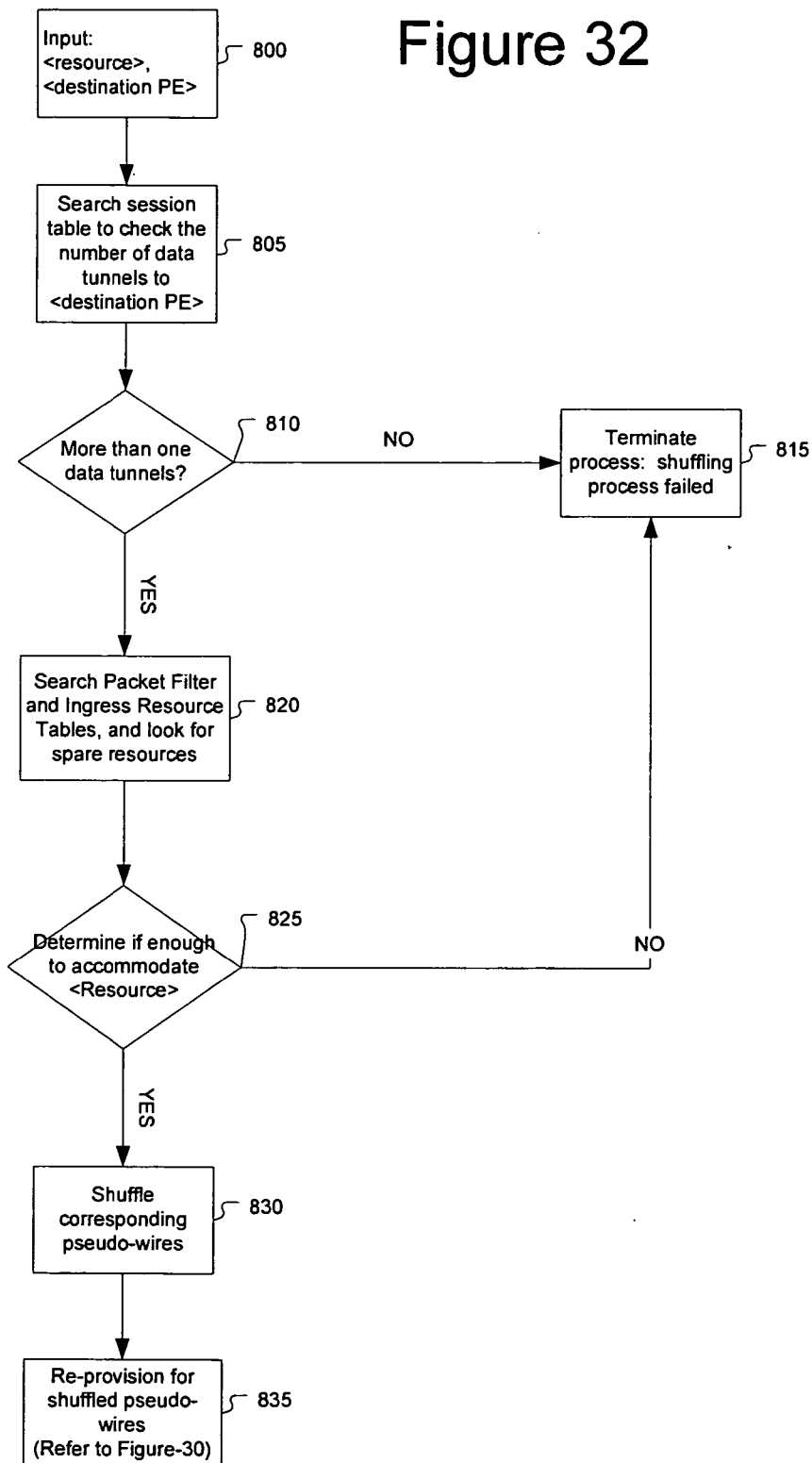
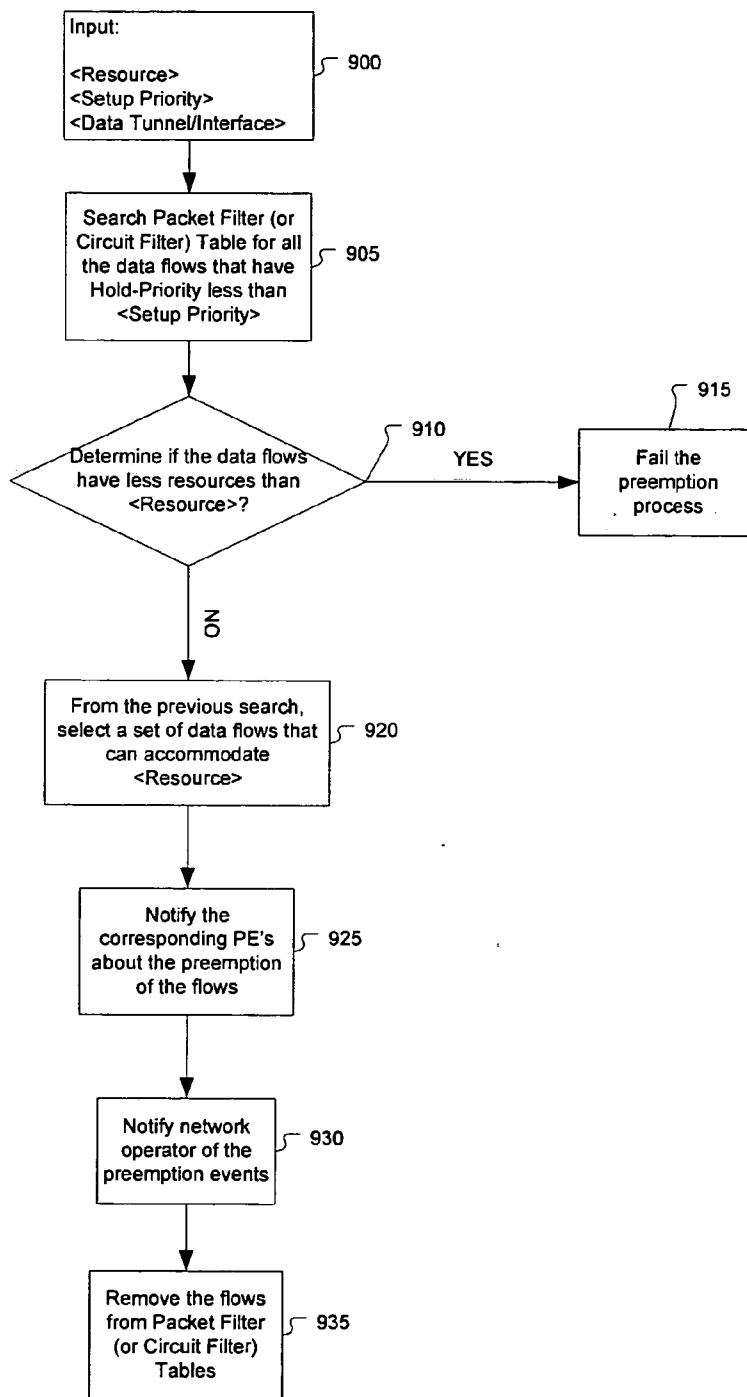




Figure 33



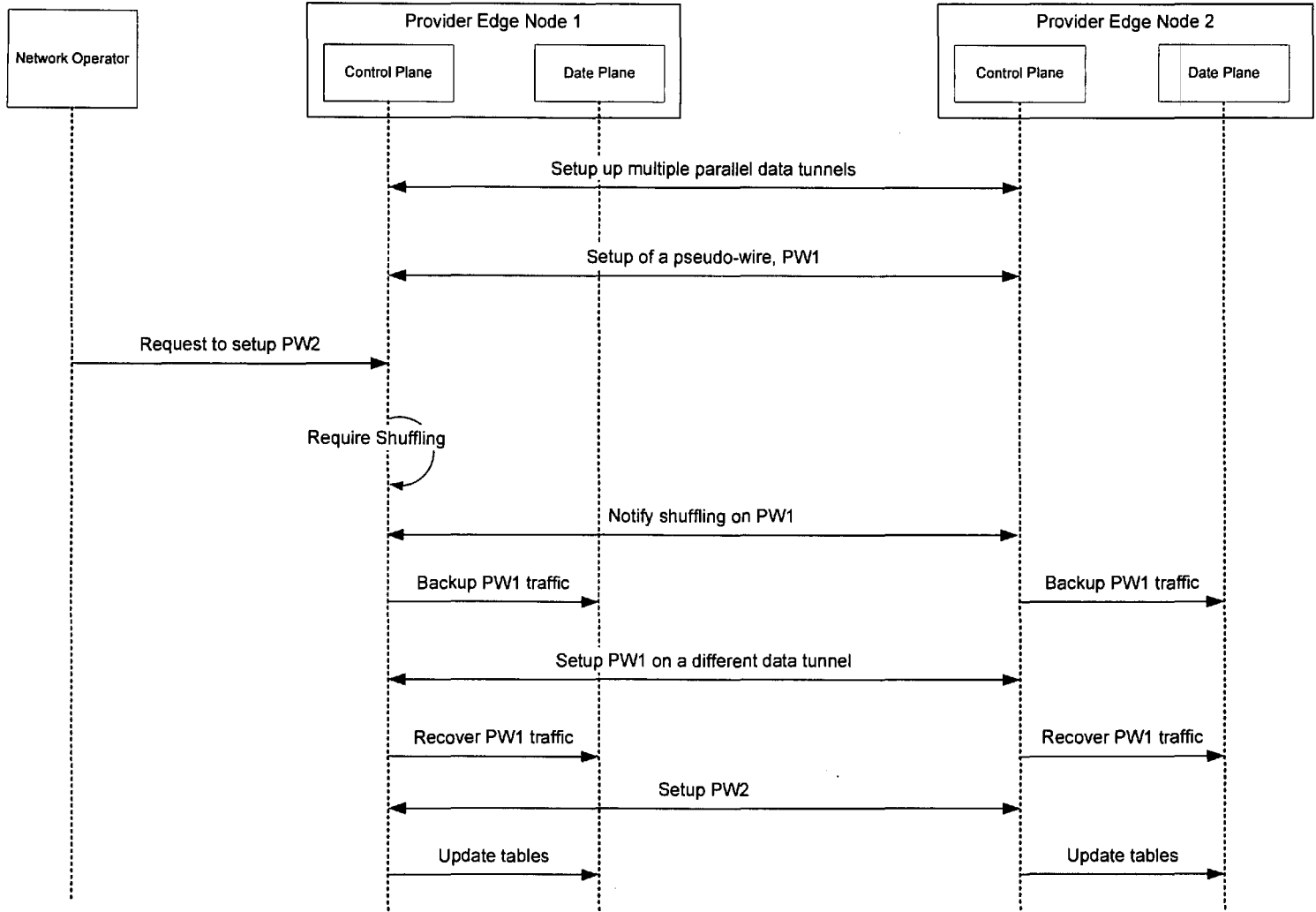


Figure 34

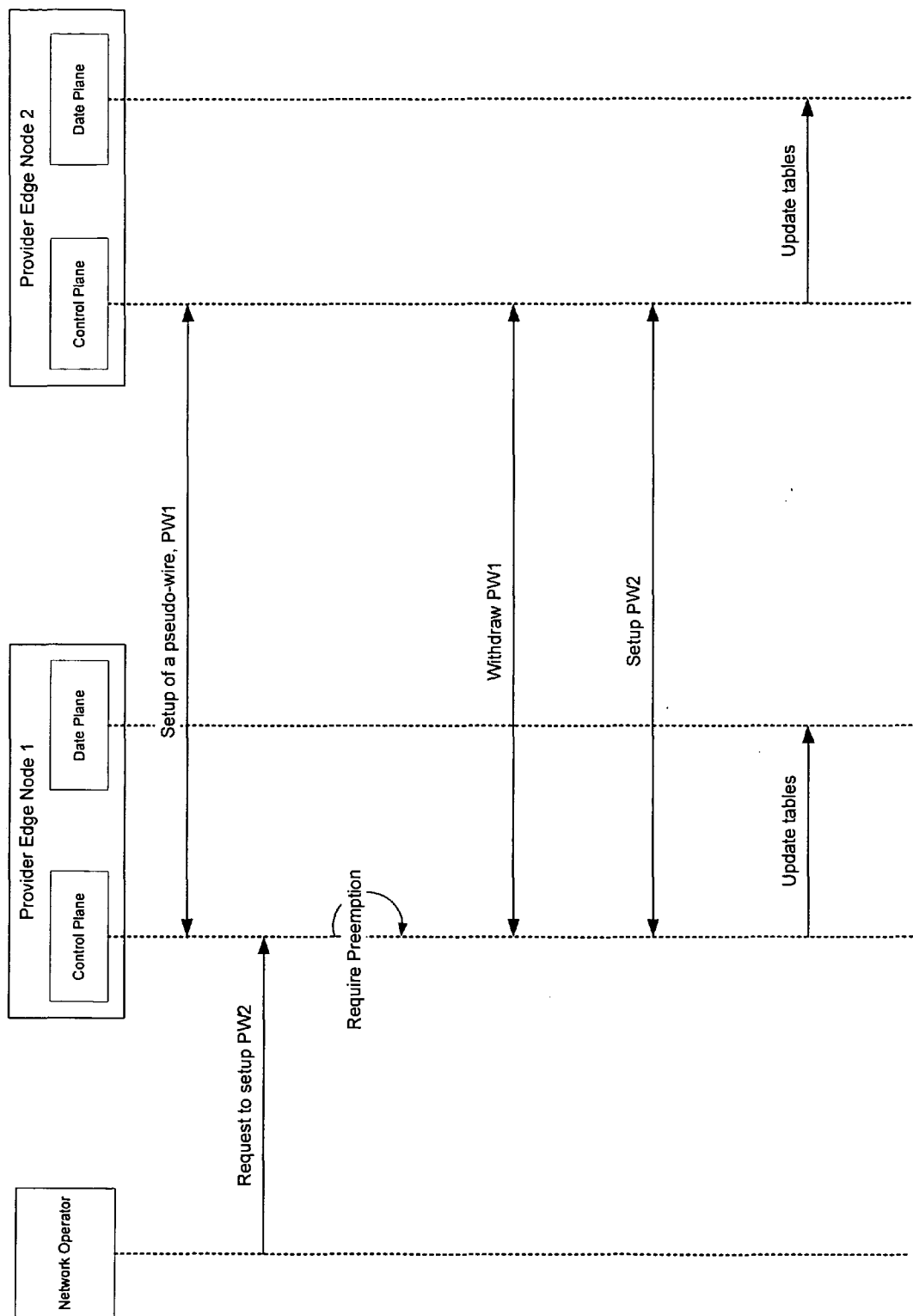


Figure 35

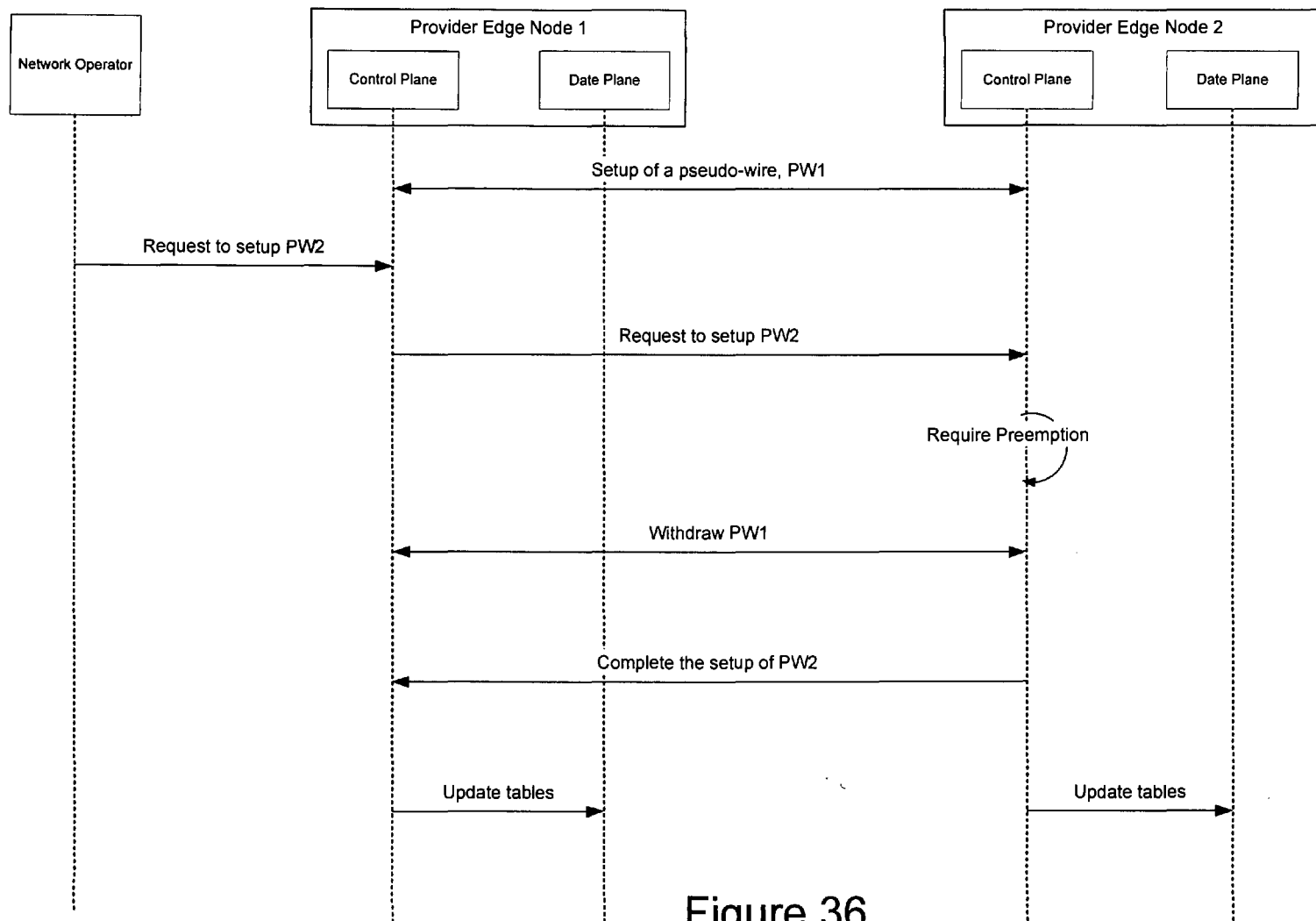


Figure 36

**METHOD AND APPARATUS FOR PERFORMING DATA FLOW INGRESS/EGRESS ADMISSION CONTROL IN A PROVIDER NETWORK**

**CROSS-REFERENCE TO RELATED APPLICATION**

[0001] This application claims priority to U.S. patent application Ser. No. 10/757,528, filed Jan. 15, 2004 (which is a provisional conversion of and claims priority to Provisional Application No. 60/440,313, filed Jan. 15, 2003); U.S. Provisional Patent Application 60/444,456, filed Feb. 3, 2003, and U.S. Provisional Patent Application 60/444,440, filed Feb. 3, 2003, all of which are by common inventors, Ping Pan and Ralph Theodore Hofmeister, all of which are hereby fully incorporated herein by reference.

**BACKGROUND OF THE INVENTION**

[0002] 1. Field of Invention

[0003] The invention generally relates to methods and apparatuses for transporting diverse traffic types such as different types of layer-2 traffic over an optical transport network such as a SONET/SDH network. The invention more particularly relates to utilizing pseudo-wires carried directly on top of the SONET, SDH, or OTN layer to transport diverse data packet traffic types such as various types of layer-2 traffic. The second embodiment of the invention expands the field of invention to also cover electrical transport networks and expands the functionality to include admission control at the ingress and egress points of a provider network.

[0004] 2. Description of Related Art

[0005] Service provider communication networks today consist of multiple types of equipment designed to transmit and route many kinds of traffic. Traditionally, these networks evolved from voice/telephone service so they were designed to carry fixed-sized circuit connections between end users. As data applications have evolved and capacity requirements have grown, several generations of packet switched networking equipment was installed into networks to route the packet data. Examples include ATM, Gigabit Ethernet, and MPLS, as shown in FIG. 21.

[0006] While new packet switching technologies continue to emerge, service providers must continue to service older technologies as it takes many years for end users to phase out a particular technology. This has led to the service providers maintaining several independent packet switched networks to carry the different types of service. Provisioning and maintaining these multiple networks is costly it would be advantageous to converge these packet switched networks onto a common network. As shown in FIG. 21, Layer-2 and MPLS switches are deployed to aggregate data flows into SONET backbone.

[0007] Conventionally circuit switched connections are used to provide transport functions between the various packet switching network equipment. But these circuit switched connections are limited in flexibility: they are available in limited bandwidth sizes: 10 Gbps, 2.5 Gbps, 622 Mbps, 155 Mbps, 53 Mbps, 1.5 Mbps, 64 Kbps, and are provisioned and maintained independently of the packet switched traffic. The static nature of these circuit connec-

tions imposes inefficiency in utilization of the capacity of the circuit switched network when carrying packet data traffic.

[0008] As a result, the interface between the packet data layer (layer 2) of the carrier network and the circuit switch layer (layer 1) leads to network utilization inefficiencies and difficult and expensive provisioning and maintenance tasks for the service providers.

[0009] The invention described herein presents a method to couple the Layer-2/MPLS packet data convergence function directly onto circuit switch equipment and integrate the control and management of connections in layer 1 and 2. Integration of these functions will greatly reduce provisioning and maintenance expenses of carrier networks and improve the utilization of the network capacity. The benefit of the invention is evident in FIG. 22.

[0010] Luca Martini and others have introduced the concept of pseudo-wires in a number of Internet Engineering Task Force (IETF) drafts, which has been widely referred to as "draft-martini". In Martini's design, some pseudo-wires can be initiated from the edge of multi-protocol label switching (MPLS) and/or IP backbone networks. Once established, a customer's layer-2 traffic can be aggregated into the pseudo-wires. To control the pseudo-wires, LDP (label distribution protocol) messages are routed through the backbone network to communicate between network edges. A serious drawback with the draft-martini design is that communication carriers must rely on MPLS/IP backbones with expensive high-performance routers to support the control messaging and label distribution protocol thereby greatly increasing the cost of transporting Layer-2 traffic which is otherwise inexpensive and relatively simple. In reality, these routers are essentially used to perform relatively trivial switching functionality.

[0011] In a parallel development, the Optical Internet-working Forum (OIF) has defined a user-network interface (UNI) specification that allows users to request the creation of Synchronous Optical Network (SONET) connections for data traffic. However, there are a number of issues in the UNI approach:

[0012] Both user and network elements must implement the UNI specification thereby dramatically increasing the cost of implementation and creating compatibility problems with non-UNI networks that interface with the UNI-enabled network.

[0013] The existing OIF UNI is only designed to interface user and network elements over optical interfaces.

[0014] George Swallow and others have proposed an overlay model where MPLS routers can use an RSVP (resource reservation protocol extension for traffic engineering) protocol to communicate with a GMPLS-enabled (generalized multi-protocol label switching-enabled) optical backbone. This approach can potentially introduce user traffic aggregation from optical network edges. However, this model requires MPLS and IP to be used across the transport networks. Also, this approach may require the carriers to reveal internal routing and resource information to the external customers, which is not practical in most of the operational networks today.

[0015] There have been a number of advancements of SONET/SDH technology in recent years. For example,

Virtual Concatenation provides the flexibility that allows edge switches to create SONET/SDH connections with finer granularity bandwidth. Link Capacity Adjustment Scheme (LCAS) uses several control bits in the SONET/SDH frame to increase or decrease a connection's bandwidth. Finally, Generic Framing Procedure (GFP) specifies the framing format for a number of link protocols, such as Ethernet and PPP.

[0016] It is admitted that MPLS, LDP, draft-martini, and OIF UNI, Virtual Concatenation, LCAS and GFP are conventional elements with respect to the invention. Although the invention utilizes some of these conventional elements, details of which may be found in available literature, the methods and apparatuses disclosed and claimed herein differ substantially therefrom. In other words, the invention leverages such conventional technologies in unique ways to achieve a method and apparatus for transporting packet data from customer data nodes over an optical network.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0017] The present invention will become more fully understood from the detailed description given herein below and the accompanying drawings which are given by way of illustration only, and thus are not limitative of the present invention, and wherein:

[0018] FIG. 1 is a network protocol layer model according to the concepts of the invention;

[0019] FIG. 2 is a simplified network diagram showing a very high level view of the inventive pseudo-wire directly over optical transport network connection techniques according to the invention;

[0020] FIG. 3 is a network operation model in a high-level block diagram format for explaining network operation according to the invention;

[0021] FIG. 4 is a structural block diagram illustrating a packet-data-enabled optical connection switch according to the concepts of the invention;

[0022] FIG. 5 is a functional block diagram illustrating operational details of the inventive packet-data-enabled optical connection switch according to the invention and further illustrating the processing of the data packets on the ingress pathway through the switch;

[0023] FIG. 6 is a functional block diagram illustrating operational details of the inventive packet-data-enabled optical connection switch according to the invention and further illustrating the processing of the data packets on the egress pathway through the switch;

[0024] FIG. 7 is a network diagram explaining the operation of control messages according to the concepts of the invention;

[0025] FIG. 7a is a detailed block diagram illustrating the structure and function of the packet processing engine according to the invention;

[0026] FIG. 7b is a diagram of the packet filter table structure according to the invention;

[0027] FIG. 7c is a diagram of the circuit filter table structure according to the invention;

[0028] FIG. 7d is a diagram of the session table structure according to the invention;

[0029] FIG. 8 is a high-level block diagram of a packet-data-enabled optical connection switch according to the invention;

[0030] FIG. 9 is a detailed block diagram of alternative construction and operation of a packet data-enabled optical connection switch and further illustrating an alternative connection of a packet access line module (PALM) according to the invention;

[0031] FIG. 10 is a high-level block diagram showing an alternative packet-data-enabled optical connection configuration according to the invention and utilizing the alternative packet access line module of FIG. 9;

[0032] FIG. 11 is a high-level block diagram showing one alternative data flow within the packet-data enabled optical connection switch configuration of FIG. 10;

[0033] FIG. 12 is a high-level block diagram showing a second alternative data flow within the packet-data-enabled optical connection switch configuration of FIG. 10;

[0034] FIG. 13 is a high-level block diagram showing a third alternative data flow within the packet-data-enabled optical connection switch configuration of FIG. 10;

[0035] FIG. 14 is a high level flowchart illustrating the general operation of the invention from both the transmit and receive perspectives.

[0036] FIG. 15a is a flow chart illustrating the inventive processing of a data packet received from a data port;

[0037] FIG. 15b is a flow chart illustrating the inventive processing of a packet fetched from an optical connection including the processing of both data packets and control messages;

[0038] FIG. 16 is a flow chart illustrating the inventive method of injecting a control message into an optical interface;

[0039] FIG. 17 is a sequence diagram showing the inventive method of setting up data flow over an optical connection;

[0040] FIG. 18 is a sequence diagram showing the inventive method of removing a data flow over an optical connection;

[0041] FIG. 19 is a sequence diagram showing the inventive method of handling the situation in which the optical connection has failed or become deactivated;

[0042] FIG. 20 is an example of the inventive apparatus and methods in operation;

[0043] FIG. 21 is a model of a conventional network used by communication providers;

[0044] FIG. 22 is a model of the network according to the principles of the invention;

[0045] FIG. 23a is a high-level network diagram showing a scenario that exemplifies the need for admission control on outgoing data flows according to the invention;

[0046] FIG. 23b is another high-level network diagram illustrating the operation of service negotiation between Provider Edge Nodes according to the concepts of the invention;

[0047] FIG. 24 is a high-level block diagram of a packet access line module and control module for performing admission control according to the invention;

[0048] FIG. 25a is a high-level flowchart illustrating the processes and methods performed by the invention for pseudo-wire admission control provisioning from the perspective of an initiating point;

[0049] FIG. 25b is a high-level flowchart illustrating the processes and methods performed by the invention for pseudo-wire admission control provisioning from the perspective of an initiating point;

[0050] FIG. 26 is a diagram of the packet filter table structure according to a second embodiment of the invention;

[0051] FIG. 27 is a diagram of the circuit filter table structure according to a second embodiment of the invention;

[0052] FIG. 28 is a diagram of the session filter table structure according to a second embodiment of the invention;

[0053] FIG. 29a is a diagram of the ingress resource table structure according to a second embodiment of the invention;

[0054] FIG. 29b is a diagram of the egress resource table structure according to a second embodiment of the invention;

[0055] FIG. 30 is a high-level flowchart illustrating the processes and methods performed by the invention for pseudo-wire admission control provisioning at both pseudo-wire initiating and terminating points;

[0056] FIG. 31 diagrammatically illustrates the concept of pseudo-wire shuffling according to the invention;

[0057] FIG. 32 is a high-level flowchart illustrating the process for shuffling pseudo-wires according to the invention;

[0058] FIG. 33 is a high-level flowchart illustrating the process for preempting pseudo-wires according to the invention;

[0059] FIG. 34 shows the operational sequence of pseudo-wire shuffling between two PE nodes in terms of both the data plane and control plane;

[0060] FIG. 35 shows the operational sequence for pseudo-wire preemption at an ingress point; and

[0061] FIG. 36 shows the operational sequence for pseudo-wire preemption at an egress point.

#### DETAILED DESCRIPTION OF INVENTION

[0062] The following detailed description of the invention refers to the accompanying drawings. The same reference numbers in different drawings identify the same or similar elements. Also, the following detailed description does not

limit the invention. Instead, the scope of the invention is defined by the appended claims and equivalents thereof.

[0063] The expression “optically communicates” as used herein refers to any connection, coupling, link or the like by which optical signals carried by one optical system element are imparted to the “communicating” element. Such “optically communicating” devices are not necessarily directly connected to one another and may be separated by intermediate optical components or devices. Likewise, the expressions “connection” and “operative connection” as used herein are relative terms and do not require a direct physical connection.

[0064] Definitions:

[0065] The invention described below utilizes various terms that may or may not be fully consistent with the conventional understanding of similar or identical terms. To clarify the meaning of these various terms the following definitions are used by this invention description:

[0066] a) MAC: media access control: The interface to the physical media. Assembles and disassembles frames and controls physical interface communications. The physical interface and frame format is L2-specific so that different client interfaces will contain specific MAC devices and/or multi-purpose MAC devices.

[0067] b) PALM: Packet Access Line Module. unit that originates and terminates packet data traffic from/to other equipment via physical interfaces. The PALM differs from the TDM (Time Division Multiplexed) Line Module in that it terminates packet data physical interfaces and frames and processes the packet traffic. The PALM described in more detail below generally contains the PPE (packet processing engine) and PPE controller and performs the translation and aggregation of packet data to/from optical connections. The simplified PALM' in the server architecture does not originate/terminate the optical connections. Instead, it translates packet data to/from internal connections between the PALM' and the server cards.

[0068] c) PPE: Packet Processing Engine. Performs per-packet forwarding decisions, appends/removes encapsulation labels, processes and delivers control messages, collects performance statistics, polices incoming traffic and shapes outgoing traffic.

[0069] d) optical circuit switch: A network element that switches and manages optical connections.

[0070] e) line module: a field-replaceable unit of the switch that contains the physical ports for traffic termination and origination.

[0071] f) data flow: a sequence of data packets that are associated with one another. All packets in a flow originate at the same node and terminate at the same node but not all packets with the same origin and termination are necessarily in the same flow as one another.

[0072] i) customer data flow: includes all types of L2 and MPLS packets. Flows from/to the client edge are differentiated by one another by various means, depending on the physical interface and frame format of the data link layer.

[0073] ii) provider data flow: also feeds into the line modules being used as well as the various node

definitions below. The invention does not depend on the topology or protection scheme of the optical network. The invention simply requires a point-to-point connection between two provider edge nodes.

[0074] g) provider edge node: the nodes at which client data packets from a flow are translated from/to an optical connection. Packets in a flow will traverse two and only two provider edge nodes: the ingress and the egress.

[0075] h) customer edge node: the node originating (terminating) the data link layer session terminating (originating) on the provider edge node client port.

[0076] i) intermediate provider nodes: nodes that the optical connection traffic passes through between the ingress and egress provider edge nodes. The intermediate nodes do not have to be aware of the data flows contained within the optical connections. Their primary function is to switch/manage the optical connection as they would a traditional or non-data flow optical connection.

[0077] j) encapsulation label: A unique identifier contained in every data packet traversing the optical connection, used to differentiate pseudo-wires. The encapsulation label is normally appended by the ingress provider edge node and removed by the egress provider edge node. However, it is possible and may be desirable in some cases for the encapsulation label to be appended and/or removed by a customer node, or over-written by an intermediate provider node.

[0078] k) pseudo-wire: a logical point-to-point connection between two provider edge nodes that is used to forward data packets from one and only one flow. One or more pseudo wires may be contained in an optical connection. A pseudo wire differs from a flow in that: 1) it originates and terminates on provider edge nodes while a flow does so on customer nodes; 2) the arrival sequence of packets will be maintained over the pseudo wire while a flow may not guarantee the sequence of packets.

[0079] l) control message label: a unique label such as the IP4 Explicit NULL Label that distinguishes control messages from data packets. In general, a unique encapsulation label to differentiate packets in an optical connection that are used by the provider nodes to pass management and control information between themselves.

[0080] m) control message: a message or signal that is used to control the provider network, customer edge nodes, components thereof, or the data being transported across the provider network or to the customer edge nodes. The invention does not generate the control messages or effect control based on them. Instead, the invention is concerned with transporting such conventional control messages. In general, the invention can practically tunnel any appropriate control message. Some illustrative but non-limiting examples are as follows:

[0081] 1. control messages relating to MPLS/IP control protocols: such control messages are used to discover and establish pseudo-wires as well as MPLS labeled-switch-path. All of these MPLS/IP control messages may be aggregated into an optical connection with label Explicit-Null or other control message encapsulation label according to the invention as discussed in detail below. Some of the more important categories of control messages may be

taken from the following protocols: LDP (label distribution protocol), RSVP (resource reservation protocol), and OSPF (open shortest path first).

[0082] 2. IP Data control messages: To ensure the connectivity between two edge nodes, the user can aggregate probing packets from an edge node, and check if they can be received at the other edge node. Such probing packets are defined in ICMP (internet control message protocol) and LSP-ping (a special sequence of packets designed to detect the connectivity of MPLS LSPs as known in the art.

[0083] 3. Layer-2 messages: To interconnect two layer-2 data interfaces through an optical connection, it is possible to tunnel conventional Layer-2 control messages such as ARP (address resolution protocol) PAUSE (a signaling protocol in Ethernet for flow control), heartbeat messages between two nodes through an identifiable control message encapsulation label according to the teachings of the invention.

[0084] 4. Control messages relating to upper application data: when supporting IP encapsulated packets, such as real-time traffic using RTP (real time protocol) which are used to convert real-time streams into IP packets. The invention can pick out or capture the in-band control packets within RTP packets such as RTCP (Real Time Control Protocol) packets and deliver them to the other edge of the optical connection. This will allow the edge nodes to monitor real-time flows, and enforce associated QoS for the flows.

#### [0085] General Description

[0086] In general terms, the invention initiates and maintains pseudo-wires directly over existing SONET networks using the already-deployed SONET switching gear. In the invention, unlike a UNI-based network, the switching intelligence only needs to be implemented in the SONET switches (network elements) and the users are not required to implement additional functionality. Furthermore, the invention works over a wide variety of customer interfaces including Ethernet, ATM, and Frame Relay optical and/or copper interfaces.

[0087] By examining some of the existing communication backbone topologies and traffic patterns, the inventors noticed that much of the data traffic comes from traditional switching networks: Ethernet, Frame Relay and ATM. Typically, voice traffic can be transported via Frame Relay circuits, and ADSL is based on ATM. With the recent rapid advancement in Gigabit Ethernet technology, Ethernet interfaces have been gradually deployed at places where both IP and non-IP traffic aggregation takes place.

[0088] Hence, the invention represents a very practical application that enables carriers to "tunnel" user traffic through well-provisioned SONET transport backbones from the edge of their networks. Further, the idea of developing yet another layer of tunnels on top of SONET cross-connections, such as building MPLS LSPs (label switched paths) as is being proposed by router vendors, is not economically practical or technically beneficial.

[0089] The invention creates "pseudo-wires" over, for example, SONET cross-connections directly, and switches



layer-2 MAC frames from network edges, reducing cost and complexity of the network switching elements. The invention may utilize many of the conventional mechanisms for setting up pseudo-wires but in unique ways as explained herein. Details of the conventional pseudo-wire mechanisms are well known and need not be discussed here in detail. Instead, this disclosure focuses on the adaptation of pseudo-wire techniques such that a pseudo wire may be carried directly over a provisioned SONET network. Alternatively, the pseudo wire may be carried directly over a provisioned Synchronous Digital Hierarchy (SDH) or Optical Transport Network (OTN) network.

[0090] The inventive protocol-layering model is shown in FIG. 1. It is important to realize that this protocol-layering model is different from the current framework, where pseudo-wires are created on top of either MPLS or IP GRE (generic routing encapsulation) tunnels which are, in turn, carried on top of the SONET transport layer.

[0091] One constraint in the conventional framework is that to create and manage MPLS or GRE tunnels (generic routing encapsulation), IP routing, IGP (interior gateway protocols) and BGP (border gateway protocol) and signaling RSVP-TE (resource reservation protocol extension for traffic engineering) and LDP (label distribution protocol) have to be used throughout the network. Therefore, to transfer layer-2 traffic according to conventional schemes such as those proposed by Luca Martini, the carriers have to rely on an IP overlay network between the layer-2 switching networks and the transport networks. Due to backbone traffic volume, high-end expensive backbone routers are required to construct such overlay networks. This design could result in adding tremendous cost to carriers, while their existing SONET transport links and equipment may be under-utilized. Also, maintaining an additional overlay IP network increases the network management and operation cost to the carriers.

[0092] Thus, to achieve the objectives of transporting layer-2 traffic, the inventors create pseudo-wires over, for example, SONET cross-connections directly, and support draft-martini (or equivalent) on SONET switches at network edges to setup and manage pseudo-wires. No router over-layer network is required in the inventive design.

[0093] Returning to FIG. 1, the protocol-layering model includes the conventional SONET transport layer that creates and maintains SONET cross connections in the conventional fashion. The pseudo-wiring is carried directly on top of the SONET transport layer according to the inventive protocol-layering model. Such pseudo-wires may be used to carry Layer-2 traffic such as Ethernet MAC, ATM, Frame Relay, etc. as well as MPLS data packets. In general, any packetized traffic may be carried by the pseudo-wires. The next layer is the actual payload which may be any data including voice, data packets, etc as is well known in the art.

[0094] It is important to realize here that, in the conventional model proposed in IETF and Luca Martini, the pseudo-wiring layer situates above IP layer. Below the IP layer is MPLS, Layer-2 and transport layers, respectively. One of the main reasons for such a model is to use IP layer for control message delivery. Since only routers have the ability to deliver control messages through the Internet backbone, pseudo-wiring therefore becomes a router-only application. In contrast, the invention utilizes the conven-

tional SONET transport layer to deliver control messages between edge nodes. As a result, pseudo-wiring can be accomplished on devices other than routers at a much cheaper cost.

#### [0095] Overview Of Operation

[0096] Before proceeding to the apparatus details, a general overview of the inventive operation is provided. Setting up pseudo wires (PW) may follow a procedure as defined in [PWE3-CTRL (L. Martini, et al, "Transport of Layer 2 Frame Over MPLS", draft-ietf-pwe3-control-protocol-05.txt)], but this procedure is modified by the invention to operate in the context of PW directly on top of the SONET, SDH, OTN or equivalent layer. The operation reference model for a SONET system is shown in FIG. 2 but it is to be understood that substantially the same reference model applies for SDH and OTN.

[0097] As shown in FIG. 2, the inventive network includes customer data nodes such as customer data nodes 1 and 2. A customer data node may be a conventional switch or router. The provider edge node generally includes conventional SONET cross-connect functionality but implemented by a data-enabled SONET switch according to the invention such as the one illustrated in FIG. 4 and further explained below.

[0098] From the customer network edge (customer data nodes as illustrated in FIG. 2 represent the customer network edge), data flow such as layer-2 frames enter the provider's backbone. More specifically, a data packet such as a layer-2 frame may be sent from customer data node 1 to provider edge node 1. The provider edge nodes 1, 2 set up a SONET cross-connection in the usual and conventional fashion across the provider network.

[0099] The invention then sets up a pseudo wire directly within the SONET cross-connect as further illustrated in FIG. 2. The pseudo-wire and SONET cross-connect are terminated at the other end of the provider network, in this case at provider edge node 2. The provider edge node 2 then transmits the data flow (e.g. layer-2 frames) to the customer data node 2.

[0100] It is to be understood that the provider network typically includes far more than 2 edge nodes and that intermediate nodes are also typically included but for ease of illustration such additional nodes are omitted from FIG. 2.

[0101] Each of the layer-2 frames within the layer 2 flow has a "flow-id" in their header. The flow-id may be implemented with conventional VLAN-ID's for Ethernet MAC frames, DLCI for Frame Relay frames, and VPI/VCI for ATM cells. It is also a possibility that the customer edge equipment may inject MPLS frames into the backbone. The use of this flow-id for the setting up and maintenance of pseudo wires according to the invention is further explained below.

[0102] FIG. 3 is a network operation model according to the invention and is useful for illustrating the general concepts of the invention. The customer data nodes (A, E) and provider edge nodes (B, D) may be implemented as discussed above. The backbone network is a conventional optical network such as a SONET, SDH or OTN-based network that is typically part of a provider network.

[0103] In reference to FIG. 3, data packets travel from A to E through B, C and D. Each packet is encapsulated with either Layer-2 and/or MPLS label. Each Layer-2 and MPLS label uniquely identifies one data flow between two nodes. In this description, when such a data flow is placed onto the provider network according to the inventive teachings it is referred to as “pseudo-wire”. Further, it is assumed that the data flows and pseudo-wires are bidirectional, although the mechanism defined here does not exclude the operation for uni-directional traffic. It is further assumed that the backbone network, C, is a conventional carrier’s transport network utilizing conventional mechanisms such as SONET-switching to deliver data. In other words, no modifications are necessary for the backbone network C elements to carry the inventive pseudo-wires.

[0104] Provider edge nodes B and D are the devices to which this invention will apply and represent the network elements that would be modified (or replaced) according to the invention. Provider edge nodes B and D are capable of performing both data switching and circuit switching. “Data switching” means that the packets are forwarded based on Layer-2 and MPLS headers. “Circuit switching” means all data sent to the circuit will be routed through the network along the same path from the time the circuit is established until it is terminated.

[0105] Upon the completion of inspecting an incoming data packet, provider edge nodes B and D will encapsulate the data packet with a label that can uniquely identify the user flow to which the packet belongs, and send the packet to a pre-established circuit over backbone network C. At egress, provider edge nodes B and D will recover the packet from the circuit, remove the label and transmit the packet out to the proper destination. There exist one or multiple circuits between provider edge nodes B and D. Each circuit can aggregate one or multiple pseudo-wires.

[0106] From the control plane perspective, it takes two steps to initiate a pseudo-wire over a circuit between provider edge nodes B and D. The first step requires the network operator, F, to download the mapping between the pseudo-wires and the circuits to the provider edge nodes B and D. The creation of the mappings may be the result of a prior business agreement, or bilateral agreement between carriers, and is beyond the scope of this invention.

[0107] Once the mapping information has been received and processed on provider edge nodes B and D, B and D will start to negotiate with each other to agree upon the encapsulation labels that pseudo-wires should use for packet encapsulation. By default, provider edge nodes B and D will allocate two encapsulation labels for each pseudo-wire, one for receiving and another for transmitting. Upon the completion of the label negotiation, provider edge nodes B and D will update the encapsulation label information to the data plane, and thus a pseudo-wire has been created.

[0108] At any given time, provider edge nodes B and D may inform operation status to network operator F. Likewise, network operator F may query provider edge nodes B and D for control and accounting information. However, it is beyond the scope of this invention to further specify the relationship between network operator F and customer (client) data nodes, A and E.

[0109] The apparatus elements within the provider edge nodes that is responsible for the functionality described

above is shown in block diagram form in FIG. 4. As shown therein, the inventive modifications are within an optical circuit switch such as a SONET, SDH or OTN optical circuit switch and transform the conventional optical circuit switch into what is termed herein a “packet-data-enabled optical connection switch” which is represented as element 5 in the drawings.

[0110] As shown in FIG. 4, the packet-data-enabled optical connection switch 5 includes a packet access line module (PALM) 10 that receives packet data from a port. This is diagrammatically indicated by a data flow arrow but the physical port will also include an appropriate physical interface (not shown) the conventional construction of which will vary depending upon the type of packet data being received and physical interface (optical, copper, line rate) as is known in the art. The PALM 10 is operatively connected to a TDM switch fabric 30 which may be constructed with a known cross connecting TDM switch fabric such as those used in conventional SONET switches one example of which is used by the CoreDirector® switch made by CIENA Corporation.

[0111] FIG. 4 is a simplified drawing for the purposes of explaining the processing of a single data flow and therefore shows only a single PALM 10 having only one port receiving a data flow. Likewise, the simplified drawing of FIG. 4 only shows one output from the TDM switch fabric to a single TDM line module 40. It is to be understood that the actual implementation would have a plurality of ports for the PALM 10. Furthermore, the actual implementation would have a plurality of ports that feed into the TDM switch fabric 30 and that the TDM switch fabric 30 output will feed into a plurality of TDM line modules 40 and output ports. In addition, it is to be understood that the implementation would have a mechanism to aggregate packet data from a plurality of PALMs prior to transmitting into the optical connections. Such a mechanism for aggregating packet data is known in the packet data switch art and could be included in the inventive packet-data-enabled optical connection switch 5, 5'.

[0112] In general, the packet fabric 34 provides connectivity between PPEs and the PPEs perform the aggregation. Even without aggregation over multiple PALMs there could still be other types of aggregation performed by the invention because a single PALM 10 may have multiple physical ports and flows from different ports may be mapped to pseudo-wires that reside in a common optical connection.

[0113] Examples of a full packet-data enabled optical switch are explained below in reference to FIGS. 8 and 10.

[0114] The TDM line module(s) 40 are conventional elements in and of themselves and provide the functions of framing (via conventional framer 45 included therein) and, electrical-to-optical conversion, and optical signal transmitting such that the data may be carried as an optical signal over the provider network. The framer 45 is a very conventional element and may utilize conventional optical transport framing schemes such as SONET, SDH, OTN, or a future developed optical signal transport framing scheme. It is greatly preferred that standardized optical transport framing schemes be used so as to take advantage of and otherwise leverage the existing optical networks utilizing such standardized framing schemes. In the U.S., this would mean SONET while in Europe it would be SDH since those are the respectively prevailing standards at this time.

[0115] The PALM 10 includes a media access controller (MAC) 12 which is a conventional element receiving packet data and terminating the customer data flow. The MAC also extracts the packet data such as an L2 packet from the customer data flow. The MAC 12 is connected to a packet processing engine (PPE) 15 that is a unique element constructed according to the principles of the invention as further discussed below in relation to FIGS. 7a-d.

[0116] The packet processing engine 15 has access to a mapping database 19-1 that contains mapping tables (packet filter table 60 subset and circuit filter table 80 subset) which are explained below in relation to FIGS. 7a-c). The PPE 15 also has access to a control message database 18-1 which includes a session table 25 subset. Generally speaking, the PPE 15 classifies the incoming packet or otherwise determines what type of packet is incoming, polices the data flow, collects performance statistics, appends an appropriate encapsulation label, aggregates traffic and shapes the outgoing traffic for logical circuits. Aggregation of traffic is possible since a single optical connection (e.g. a subnetwork connection which may be at a rate of OC-12, OC-48, etc) may hold more than one pseudo wire containing a packet. Further details of the PPE 15 operation are provided below in relation to FIG. 7a.

[0117] The PPE is operatively connected to the mapping engine 17 which is itself a conventional element that encapsulates the packet+label. One example of such encapsulation that may be used by the invention is the conventional GFP (Generic Framing Procedure as defined by ITU-T G.7041/Y.1303). Other examples include LAPS (Link Access Procedure-SDH, ITU standard X.86), PoS (Packet over SONET IETF RFC2615) and other HDLC-framing methods (such as the ones used on Cisco routers).

[0118] The mapping engine 17 also originates and terminates optical connections as is known in the art (e.g. optical connections using SONET, SDH or OTN). The mapping engine, in one implementation originates/terminates the optical connection. The TDM fabric 30 and TDM LM/framer 45 allow muxing/demuxing of the optical connection so that it may go out one or more physical ports and share the physical port with other TDM traffic and/or other PW-carrying optical connections. These optical connections output from the mapping engine are then sent to the TDM switch fabric 30 that switches the connections (or circuit elements if virtual concatenation is used). The switch fabric 30 is connected to a TDM line module 40 which includes a framer 45 that implements a conventional SONET, SDH, or OTN optical transport framing and originates/terminates the optical transport signal to/from the provider network.

[0119] As mentioned above and as shown in FIG. 4, the main data flow pathway through the packet-data-enabled optical circuit switch 5 is a bidirectional pathway. Although the above description mainly focuses on the left-to-right (ingress) flow taking the customer data flow and processing it to output an optical signal on the provider network, the reverse (egress) flow is also part of the invention. This is further discussed below in relation to, for example, FIGS. 5 (ingress flow) and 6 (egress flow).

[0120] As further shown in FIG. 4, a switch controller 20 has control over the MAC 12, PPE 15, mapping engine 17, TDM switch fabric 30 and TDM line module 40. The switch controller 20 may be constructed with, for example, a

general-purpose microprocessor and associated memory, ASIC(s), FPGA(s) or other well-known techniques for building such control modules. The control functions performed by the switch controller 20 are programmed into the microprocessor, ASIC, FPGA, etc. Conventional aspects of switch controller 20 functionality such as certain conventional aspects of control over the TDM switch fabric 30, line module 40, MAC 12 and mapping engine 17 are not described in detail herein. As appropriate, this disclosure focuses on the novel aspects of control exercised by the switch controller 20 and are explained in detail below particularly in relation to FIGS. 17-19. Generally speaking, the PW label negotiation is performed by the switch controller 20 as the PPE 15 typically cannot provide system-wide label allocation and network view etc. Once the labels have been negotiated, the switch controller 20 will download the negotiated labels to the PPEs 15 for data switching. The switch controller 20 has access to a control message database (DB) 18 which includes a session table 25. The database 18 holding session table 25 may be stored in a separate memory module as shown in FIG. 4 or it may be stored in a common memory module along with the mapping tables of database 19. More specifically, the switch controller 20 maintains a master copy of all information including a master copy of the control message database 18 storing the session table 25 and a master copy of the mapping database 19 including the packet filter table 60 and circuit filter table 80. The switch controller 20 distributes the information from all of these tables to the PPE 15 on each individual PALM 10.

[0121] In a full packet-data-enabled optical connection switch 5 such as the one shown in FIG. 8, the switch controller 20 controls a plurality of PALMs 10-1 through 10-n each of which includes a PPE 15. Continuing this notation, the individual PPEs 15-1 through 15-n each have a corresponding subset of the control message database 18 and the mapping database 19. Thus, the individual PPEs 15-n each store a control message DB subset 18-n (storing a session table 25 subset) and a mapping DB subset 19-n (storing a packet filter table 60 subset and a circuit filter table 80 subset).

[0122] FIG. 5 further illustrates the inventive ingress processing of packet data arriving as a client signal. In detail, FIG. 5 shows the main elements of the packet-data-enabled optical connection switch 5 including MAC 12, PPE 15, mapping engine 17, TDM switch fabric 30 and framer 45. A customer data flow arriving at the MAC 12 may be in a wide variety of formats including but not limited to GE (gigabit Ethernet), LAPS (link access procedure—SDH), EoS (Ethernet over SONET), ATM (asynchronous transfer mode), FR (frame relay), RPR (Resilient Packet Ring IEEE 802.17), POS (Packet over SONET) or any layer 2 packet with or without an MPLS label.

[0123] All of these data types are represented in FIG. 5 as a layer-2 packet (L2 pkt) after the associated transport frame structure has been removed. As shown therein, the MAC 12 extracts the L2 packet. The PPE 15 appends an appropriate encapsulation label (further discussed below) which is shown as "L2 pkt/Label" in FIG. 5. The mapping engine encapsulates the L2 packet with the encapsulation label in a GFP frame or equivalent and optical connection frame structure. The mapping engine further encapsulates the packet as necessary in a compatible format for the TDM

switch fabric **30**. The packet traverses the TDM switch fabric in the optical connection to one or more framers **45** where the optical connection may be groomed with other optical connections and prepared for transmission in a conventional optical frame such as a SONET frame, SDH frame or OTN frame.

**[0124]** FIG. 6 further illustrates the reverse or egress path through the packet-data-enabled optical connection switch **5** from the perspective of the data flow through the packet-data-enabled optical connection switch **5**. Specifically, data packets transmitted through the optical transport network via pseudo-wires carried on optical connections are received by the framer **45** where the underlying SONET, SDH, or OTN frame structure is terminated and the payload envelope is converted as necessary into a compatible format for the TDM switch fabric **30**. The data packets traverse the TDM switch fabric to the mapping engine **17**, which converts as necessary from the TDM switch fabric format, terminates individual optical connections, and extracts packets and removes the GFP or equivalent frame overhead. The underlying packet that still includes the encapsulation label is passed to the PPE **15**. The PPE determines the appropriate physical port to send the packet out on and optionally overwrites the L2 label based on the encapsulation label value and the optical connection it was received on. The PPE removes the encapsulation label prior to passing the L2 packet to the MAC. The MAC encapsulates the L2 packet in the appropriate L1 frame/format and sends it to the physical port for transmission to the customer edge node.

**[0125]** Edge-to-Edge Message Tunneling

**[0126]** FIG. 7 illustrates the structure of a network that can aggregate multiple data flows over a single optical connection. There exists an optical connection between two Packet-Data-Enabled Optical Connection Switches, C and H that are built according to the invention (e.g. the packet-data-enabled optical connection switch **5**, **5'** as described herein). The remainder of the nodes A, B, D-G, I and J are conventional equipment. The optical connection can be in the form of, for example, a SONET, SDH or OTN transport circuit.

**[0127]** The optical connection can aggregate multiple data flows from Customer Nodes A, B, I, and J. Each flow is associated with a unique encapsulation label at either receive or transmit direction. The packets that belong to a particular flow will be encapsulated with a label at C and H. The value of the label is the result of control-plane negotiation between C and H as further explained below.

**[0128]** One critical issue in this architecture is the delivery of the control messages. Obviously, to support large number of data flows, each Data-Enabled Optical Switch may require processing a large volume of control traffic. There are a number of methods to accomplish this including:

**[0129]** 1. Route control messages through the network. This is the method used in the Internet, where each control packet is delivered hop-by-hop until it reaches to the final destination. Note: in the similar method of aggregating data flows over MPLS network [draft-martini], the control packets are "routed" through the router network. This approach is not practical in optical networks, since this would require every optical node to establish a special

connection to a neighboring optical node for the purpose of delivering control messages only.

**[0130]** 2. Send control messages through SONET DCC channel: the DCC channel is a set of control overhead fields in SONET frames. It has been used to exchange control messages between optical nodes within optical networks. DCC channels, however, have very limited bandwidth. The option of inserting data-control messages to DCC channels may cause traffic congestion which would result in optical network internal information loss.

**[0131]** 3. Out-of-band signaling: Like SS7 networks operated in PSTN networks, one option is to build an out-of-band control network for control message delivery. However, this can be very costly in terms of network manageability.

**[0132]** After evaluating all the existing options, the inventors created an in-band method for control message delivery. The idea is to treat control messages as regular data packets, and inject them into the optical connection that they are supposed to provision for the data flows. In other words, in the invention, all control packets are to be "tunneled" through SONET (or SDH or OTN) cross-connections as regular payload from the edge. Each data flow is associated with a label, and the invention encapsulates each control message with an identifiable encapsulation label that can be recognized by the edge nodes.

**[0133]** In FIG. 7, there exists an optical connection going through nodes C, D, E, G and H. The provider edge nodes D and H include a data enabled optical switch **5** according to the invention such that C and H will use the connection to exchange control messages. Each control message is encapsulated with a label that both C and H can recognize. Subsequently, C and H will capture and send the control messages to the control plane for processing. One example of an identifiable label is the Explicit NULL label defined in Rosen et. al, "MPLS Label Stack Encoding", RFC3032, Network Working Group, Request for comments 3032 submitted to Internet Society, January 2001 which may be found at <http://www.ietf.org/rfc/rfc3032.txt>. The identifiable label is also called a control message encapsulation label herein and is not limited to the NULL label mentioned above. Indeed, any label could be used as the control message encapsulation label. For example, the provider edge nodes may negotiate any label to serve as the control message encapsulation label and such a label will thereafter identify the data packet as a control message.

**[0134]** There are a number of advantages in the inventive approach described herein including:

**[0135]** 1. Control message processing only involves the edge nodes. Network intermediate nodes are not disturbed, need no modification and merely pass along optical signals in the normal fashion. In FIG. 7, other than the provider edge nodes C and H, the rest of the optical nodes (D, E, F, and G) are not aware of the existence of control messages.

**[0136]** 2. Since control messages are encapsulated with labels, this simplifies the processing overhead at the provider edge nodes. The control messages are processed as regular data packets. Instead of sending out to a data interface, they are forwarded to the

control module. The detailed mechanism for accomplishing this is elaborated upon below.

[0137] 3. Since control messages traverse the same optical connections that data flows will traverse, it is easier and faster for the edge nodes to react to network failures. In comparison, in MPLS networks, when there is a failure on the data plane, it will take seconds before the control plane will be aware of the problem—likely to be notified from the routing protocol updates. In the inventive approach, the control-plane and the data-plane share the same fate. As a result, the control-plane can respond to failures faster. This is a huge advantage particularly because protection mechanisms can be triggered much faster thereby preventing data loss. At modern line rates currently approaching 40 gigabits/seconds per wavelength activating protection mechanisms in a shorter time will prevent the loss of tremendous amounts of data.

[0138] Generally speaking, the invention operates as follows. When a data flow such as a layer-2 frame is received from a user's network, the PPE 15 encapsulates (or pushes) a pre-negotiated encapsulation label onto the packet. On the other hand, when a control packet (such as LDP Hello message) needs to be delivered through the network, the invention pushes an identifiable label such as the "IP4 Explicit NULL Label" on to the control message. The PPE 15 will direct all frames into the pre-established SONET connections (pseudo-wires). Further detailed operation is provided below in relation to FIGS. 14 and 16.

[0139] On the other end of the SONET connection, the PPE 15 will de-encapsulate (or pop) all received frames. For data packets, the PPE 15 forwards them to the user network. If the received label is the identifiable control message label (e.g. "IP4 Explicit NULL Label"), the PPE 15 forwards the message to the switch's central processor 20 for further processing. Further details are provided below in relation to FIGS. 14 and 15b.

[0140] FIG. 14 is a high level flowchart illustrating the general operation of the invention from both the transmit and receive perspectives. All of the operations outlined in FIG. 14 are performed by the PPE 15. As shown therein, the invention first establishes (300) an optical connection between two provider edge nodes which is a conventional process in and of itself that may use conventional SONET, SDH or OTN techniques to do so. The data packets may be aggregated into this optical connection. Next, the PPE 15 tunnels (305) packet data within the established optical connection. Pseudo-wires may then be established (310) by tunneling command messages within the same established optical connection as used for the packet data. Like the data packets, the control messages may also be aggregated within the same optical connection at least to the extent the control messages share the same optical connection pathway through the provider network. When transporting control messages, the PPE utilizes (320) a distinguishable encapsulation label for the command message. Such a distinguishable encapsulation label is also referred to herein as a control message encapsulation message.

[0141] On the receive end, as further shown in FIG. 14, the PPE 15 parses the encapsulation label from the received data. The PPE 15 may then decide (330) whether the parsed

encapsulation label matches the command encapsulation label type. If yes, then the received message is processed (340) as a command message a process which may include sending the command message to the switch controller 20. If the parsed label does not match the command message encapsulation label type, then the received message is processed (335) as a data packet a process which may include using the parsed label to lookup the outgoing data interface from the circuit table 80 that applies to the particular data packet just received.

[0142] FIG. 7a is a detailed block diagram of the packet processing engine (PPE) 15 that is a key part of the invention and which may, for example, be part of the packet access line module 10 as shown in FIG. 4.

[0143] The packet processing engine 15 is the device responsible for processing incoming data packets, mapping packets into optical connections, processing packets received from optical connections, and injecting control messages into optical connections. Unlike traditional switching devices that perform either packet or circuit switching, in the invention design, each PPE 15 operates for both packet and circuit switching simultaneously.

[0144] The processing of data packets includes operations such as packet header lookup, extra header encapsulation, and packet switching into optical connections. The processing of packets from optical connections includes operations such as SONET Path Over Head (POH), packet header manipulation and label switching. One SONET POH handling is the ability to work with Virtual Concatenation and LCAS that are used to group and maintain optical connections.

[0145] The PPE 15 includes a packet filter 65 receiving data packets as shown from the MAC 12. The packet filter 65 has an operative connection to packet filter tables 60 (actually a subset of all the packet filter tables as discussed above in relation to FIG. 4).

[0146] Packet filter 65 is the engine that processes the packets from data interfaces. The packet filter 65 is associated with and has access to packet filter table 60. For each incoming data packet, the packet filter 65 will extract data interface information and the packet's Layer-2 and/or MPLS headers, and use the packet filter table 60 to determine the encapsulation labels and the corresponding logical connection. The packet filter 65 forwards the packets into the corresponding optical connections so determined.

[0147] Packet filter 65 is connected to a packet forwarder 75 which is responsible for adding/stripping the labels, and forward packets to/from data and circuit interfaces.

[0148] Elements 65, 75, and 85 may be implemented any number of ways and with any number of physical elements such as logical entities within a software program. For high packet-switching performance, Elements 65, 75 and 85 can be implemented with specialize ASIC, FPGA, or off-the-shelf Network Processors. To satisfy pseudo-wire QoS requirements, further ASIC, FPGA and off-the-shelf Traffic Management chips may be required. Another example is a network processor unit complex which would include a network processing unit (NPU), memory, and optionally a traffic management chip with software coded to implement the invention running on the NPU. Another option would put all of these functions on one or more ASICs.

[0149] Packet forwarder **75** is also connected to a circuit filter **85** which has access to circuit filter table **80** (again, a subset of the circuit filter table maintained by the switch controller **20** as discussed above in relation to **FIG. 4**).

[0150] The circuit filter **85** is the engine that processes the packets coming from optical connections. Circuit filter **85** is associated with and has access to the circuit filter table **80**. For each packet fetched from the optical connection, circuit filter **85** will extract the encapsulation label that identifies the data flow from the packet, and search the circuit filter table **80** for the outgoing data interface. If the packet is a control message (as determined by the identifiable encapsulation label for control messages), it will be forwarded to the switch controller **20** via the control message pathway as further shown in **FIG. 7a**. Otherwise, the circuit filter **85** strips off the label, and forwards the recovered packet to the corresponding data interface.

[0151] PPE controller **70** has a control connection to packet forwarder **75** and a control message pathway to switch controller **60**. In addition, PPE controller **70** has access to session table **25** (again, a subset of the session filter table maintained by the switch controller **20** as discussed above in relation to **FIG. 4**).

[0152] The PPE Controller **70** is the logical entity that communicates with the switch controller **20**. PPE controller **70** is associated with and has access to the session table **25**, which maintains the mapping of control messages and outgoing optical connections. To inject a control message, PPE controller **70** searches the session table **25** to determine the encapsulating label and optical connection. Once the information is located, PPE controller **70** will encapsulate the control message and send out the control message via the optical connection (by way of the mapping engine **17**, TDM switch fabric **30**, and TDM line module **40**).

[0153] The packet filter **65** and circuit filter **85** may be constructed as logical elements within a software program. As such these filters **65**, **85** may share processing resources with the PPE controller **70** or may be separately constructed.

[0154] In more detail and as shown in **FIG. 7b**, the packet filter table **60** has the following attributes:

[0155] A Searching Key which includes the packet's (incoming) data interface and label information.

[0156] (Incoming) data interface: This is the interface that receives the packet. It can be the identification for either a physical or logical interface. The invention makes no assumption on how such information is actually obtained. However, the interface information is required for each packet being received.

[0157] Label: This can be, for example, a Layer-2 or MPLS header. A Layer-2 header can be an Ethernet MAC and VLAN tag, a Frame Relay DLCI, or an ATM VCI/VPI. It is noteworthy that a received packet may have been encapsulated with a Layer-2 header and a MPLS label. In this case, two matching keys are defined: one with Layer-2 header; the other, MPLS label. In **FIG. 7b**, Packet-Filter-1 and Packet-Filter-4 can be applied to the same packet.

[0158] Outgoing Optical Connection: This is the connection that the packet will be injected into as it enters the provider network.

[0159] Encapsulation Label: The label for each data flow. It will be encapsulated with the packet.

[0160] Filter Priority: The importance of the filter. As mentioned above, a packet may be encapsulated with both Layer-2 and MPLS. Thus, two matching filters may be found. We use the Filter Priority to decide which filter should be applied to the packet. In **FIG. 7b**, if a packet received from Port-1 that matches to both Packet-Filter-1 and Packet-Filter-4, Packet-Filter-1 will be chosen since it has a higher priority.

[0161] Guaranteed QoS: This is an optional field when QoS (quality of service) is an issue. If so, each data flow should comply within a fixed traffic boundary. Otherwise, traffic congestion may result within an optical connection. This field maintains the guaranteed QoS for the flow. For packets that do not comply, a user-defined traffic conditioning mechanism will be used. The mechanism itself is beyond the scope of this invention.

[0162] As shown in **FIG. 7b**, the packet filter table **60** is populated with data showing the various types of packets that may be processed including Ethernet, ATM, FR and MPLS. Indeed, in this populated packet filter table **60**, the PPE **15** is handling **4** different flows each with a unique encapsulation label. The corresponding outgoing optical connection fields are associated with each of these packet types.

[0163] As further shown in **FIG. 7a**, the data packets that arrive at the packet filter may be in the form of a layer 2 data (L2) with an associated packet or frame structure encapsulating the L2 data. Alternatively, an MPLS data with an associated packet or frame structure may also arrive at the packet filter **65**. At element **75**, the element **75** pushes a pre-negotiated encapsulation label onto the L2 packet or MPLS packet. When a control message is received from switch controller **50** via PPE controller **70**, the element **75** also pushes a pre-negotiated encapsulation label onto the control message. With the encapsulation label added, the data flow is next sent to the circuit filter **85** before being output as a logical circuit (SNC or sub-network connection) to the next stage which is the mapping engine **17** as shown in **FIG. 4**.

[0164] **FIG. 15a** shows in more detail the processing performed by the PPE **15** on a data packet received from a data interface. As shown therein, the PPE **15** receives (**400**) a packet from a data port and then the packet filter **65** parses (**405**) the layer-2 (and perhaps the MPLS header if present) and searches the packet filter table **60**.

[0165] The packet filter **65** then decides (**410**) whether there is a match with the packet filter table **60**. If not, then the packet is dropped (**440**) thereby ending processing for the received packet.

[0166] If there is a match, the flow proceeds and decides (**415**) if there is more than one matching filter which may be the case if the packet is encapsulated with both Layer-2 and MPLS headers (or other multiple headers as may be the

case). More than one header cases the packet filter **65** to choose (**445**) the header with the highest priority (see filter priority field in Fib. *7b*).

[**0167**] The traffic condition may then be determined (**420**). When a filter is found for a packet, the traffic condition for that flow, such as the bandwidth consumed by the flow, will be known. The packet filter **65** and packet filter table **60** keep track of the QoS information for all flows. If, by receiving this packet, it will cause the flow's QoS parameters (such as bandwidth consumption) to be over its defined limit, the PPE **15** will apply traffic conditioning to the packet, either dropping or tagging the packet. With this information, the packet filter **65** may then determine (**425**) if the traffic condition is within a QoS limit. The invention does not define the actual mechanism for the packet filter **65** to come to that decision **425**; rather, it only operates on the final outcome. If not within the QoS limit, then the traffic condition or rule is followed **450** meaning that the traffic is dropped or tagged. If (**455**) not tagged, the packet is dropped (**440**). If it is tagged, the flow proceeds to the encapsulation (**430**) step. Steps **420**, **425**, **450**, **455** are considered option and implemented only when QoS is a factor.

[**0168**] The encapsulation (**430**) involves looking up the encapsulation label from the packet filter table **60** and pushing the encapsulation label onto the packet as illustrated in **FIG. 7a**. Then, the encapsulated packet may be sent (**435**) out to the outgoing optical connection as defined in the packet filter table **60**.

[**0169**] In general, the PPE **15** performs the following processes. Since each SONET cross-connection can carry traffic from multiple L2 users, it is necessary to be able to distinguish individual user's frames at place where demultiplexing takes place. The PPE takes care of this by pushing an encapsulation label onto every L2 frame that will enter the provider network. The encapsulation label may come from the negotiation between provider edges using LDP.

[**0170**] At exiting edge, the encapsulation label will be popped, and the original frames will be recovered and delivered out to the destination customer. This process is described below in more detail in relation to **FIG. 15b** and the circuit filter table of **FIG. 7c**.

[**0171**] **FIG. 7c**: Circuit Filter Table

[**0172**] The Circuit Filter Table has the following attributes:

[**0173**] Searching Key: Optical Connection and Label

[**0174**] Optical Connection: The connection where a packet is received. It can be a SONET VCG (Virtual Concatenation Group) or an optical interface

[**0175**] Label: This is the label that has been inserted at the ingress of the data flow. It is used to identify a specific data flow.

[**0176**] Outgoing Data Interface: The interface where the packet to be forwarded. As shown in **FIG. 7c**, all control messages go to "Host Interface", which is the Switch Controller in this case.

[**0177**] Overwritten Label: It is possible that the customer may want to change a packet's Layer-2 label

as it traverses through the optical network. One such instance is that the customers want to change Ethernet VLAN values to satisfy Ethernet bridging protocol requirements. Overwritten Label contains the new label information. PPE is responsible for the label over-writing.

[**0178**] Guaranteed QoS: Each data flow must comply within a fixed traffic boundary. Otherwise, this may result in traffic congestion at outgoing data port. This field maintains the guaranteed QoS for the flow. For packets that do not comply, a user-defined traffic conditioning mechanism will be used. The mechanism itself is beyond the scope of this invention.

[**0179**] As shown in **FIG. 15b**, the PPE **15** performs the following processes when receiving a packet from an optical connection. First, the PPE fetches (**500**) the packet from an optical connection. The circuit filter **85** may then parse (**505**) the encapsulation label from the packet and use it to search the circuit filter table **80** (see **FIG. 7c**). The results of the circuit filter table **80** search are used to determine (**510**) if there is exactly one match. If not, the packet is dropped (**540**) and this event is recorded.

[**0180**] If there is only one match, then the circuit filter **85** may determine (**515**) the traffic condition. Once again, the circuit filter is keeping track of the QoS parameters, (bandwidth, delay, and packet dropped etc.) for every flow. If by receiving this packet causes the flow's QoS parameters going over the limit, we will have to either drop or tag the packet.) The results of this determination (**515**) are used to decide (**520**) if the traffic condition is over the QoS limit. If yes, then the packet is (tagged or dropped) (**545**) according to the QoS rule stored in the circuit filter table **80** for that packet. A decision (**550**) is based on whether the packet is tagged or dropped: if to be dropped the flow proceeds to drop (**540**) the packet; otherwise, the flow proceeds to remove (**525**) the encapsulation label. Like the QoS processing described above in relation to **FIG. 15a**, these steps are option if QoS is not a factor in the system.

[**0181**] After removing (**525**) the label, the circuit filter **85** decides (**530**) whether to require overwriting of the packet header. See the description for the parameter above for details. If yes, the circuit filter **85** overwrites the header according to the entry for that circuit contained in the circuit filter table **80**. If the entry indicates that the label is not to be overwritten than the PPE **15** sends out the packet through the data interface defined in the circuit filter table **80** for that packet. In this way, the data flow arriving from the provider network may be correctly routed to the correct data interface and, ultimately, to the correct client edge node.

[**0182**] Since the control messages come as labeled packets, the circuit filter table **80** will match them to "host interface". The sending step **535** will send regular packets to data interfaces, and control messages to this "host interface" which is the switch controller **20** itself.

[**0183**] **FIG. 16** and session table *7d* further explain the control messaging procedures. PPE controller **70** implements the process of **FIG. 16** with access to the session table **25** of **FIG. 7d**.

[0184] FIG. 7d: Session Table

[0185] The Session Table 25 has the following attributes:

[0186] Searching Key: Control Message ID

[0187] Each control message carries a unique ID to identify which “peering session” it belongs to. A peering session is a logical connection between two edge nodes. It is used to exchange control information between two nodes. For example, in pseudo-wire operation, the customer may apply LDP [RFC3036] to negotiate data flows. LDP operates over TCP. Between two edge nodes, all control messages go over a TCP session that can be uniquely identified with TCP Sender Port Number, and IP addresses. In this invention disclosure, we shall not specify the exact message ID format. However, it is reasonable to assume that each control message carries enough information to identify the session to which it belongs.

[0188] As an example, in FIG. 7d, there are three sessions that are identified with TCP and UDP port numbers.

[0189] Outgoing Optical Connection: This is the connection that the control messages will be injected into.

[0190] Encapsulation Label: The identifiable label for the control message. PPE will insert this label to the control message.

[0191] Guaranteed QoS: All control messages within a session will have a fixed network resource level. This is designed to protect the control messages from potential congestion caused by regular data traffic.

[0192] The process begins by the PPE controller 70 receiving (600) a control message from the switch controller 20 which is then parsed (605) to find the ID as explained above.

[0193] The PPE controller 70 then searches (610) the session table 25 according to the control message ID parsed (605) from the control message. The results of the search are used to decide (615) if there is a match such that the corresponding entry may be retrieved from the session table 25. If not match, the message is dropped (640) and the event recorded. If there is a match, the PPE controller 70 may perform some QoS processing (steps 620, 625, 645, 650, 640) that are analogous to the QoS processing described above in relation to FIGS. 15a and 15b such that a repetition here is not necessary. Again, this QoS processing is considered an optional but desirable feature.

[0194] After QoS processing, the PPE may then send (635) out the control message to the associated optical interface (identified by the entry in the session table 25 for that control message) as a data payload. Specifically, the control message is tunneled as payload within a SONET, SDH or OTN frame payload and thereby shares its fate with the packet data being carried by the provider network.

[0195] Provisioning of Pseudo-Wires

[0196] The conventional LDP (label distribution protocol, RFC3036) is used by the invention to setup and manage pseudo wires: each pseudo-wire runs over a bi-directional

cross-connection such as a SONET, SDH, or OTN cross-connection. Each pseudo-wire includes two unidirectional paths, one in each direction. Each provider edge initiates the setup of the path on behalf of ingress L2 traffic.

[0197] Each path may be uniquely identified by the triple <sender, receiver, encapsulation label>. The triple is part of the message sent between nodes during the label negotiation phase shown in FIG. 17. The VCID is an example of an encapsulation label that may be used by the invention. A conventional VCID label is a 32-bit quantity that must be unique in the context of a single LDP session between two provider edges. For a given pseudo-wire, the same encapsulation label (e.g. VCID) must be used when setting up both paths.

[0198] As described during our discussion on FIG. 3, to aggregate a data flow and thus establish a pseudo-wire, the network operator first downloads all the mapping information to the provider edge nodes. Through LDP, two provider edge nodes negotiate encapsulation label for a data flow.

[0199] To create a pseudo wire between two provider edges, the network operator needs to provide the IP addresses of the provider edges, and assign a, for example, 32-bit VCID to represent this pseudo wire. To support Ethernet VLAN services, the operator needs to feed VLAN-ID's to both provider edges as well.

[0200] Through LDP, two provider edge nodes exchange encapsulation label, physical port and VLAN information, and negotiate the encapsulation labels. Specifically, LDP will use Virtual Circuit FEC and Generic Label FEC during label negotiation. Upon completion, the provider edge nodes will program hardware for frame classification and MPLS label encapsulation. The detailed operation of LDP is conventional and beyond the scope of this invention.

[0201] FIG. 17 further explains the process of setting up a pseudo wire over optical network according to the invention. Essentially, FIG. 17 is a sequence diagram that performs the following processes.

[0202] 1. Initially, there exists an operational optical connection between provider edge nodes (Node-1 and Node-2 in FIG. 17). Traditionally, in carrier networks, such connections are static in nature—they are not frequently modified once established.

[0203] 2. Node-1 and Node-2 will establish a peering session over the optical connection. The method for session establishment is to inject control messages into the connection, and each control message is encapsulated with an identifiable label. (See the description for FIGS. 7 and 7d above)

[0204] 3. Upon the establishment of the peering session, Network Operator will issue data flow setup requests to both Node-1 and Node-2. The request will include the following information:

[0205] a. The data interfaces that packets will traverse.

[0206] b. The optical connection the packets need to aggregate into.

[0207] c. The QoS (bandwidth) requirements for each flow.



- [0208] d. Optionally, the packet Layer-2 label to be overwritten (see the description for FIG. 7c)
- [0209] 4. The integrity of the requests is maintained by the network operators, and is beyond the scope of this invention.
- [0210] 5. Node-1 and Node-2 will exchange control messages and negotiate the labels to be used by the data flows. An example of the label negotiation is described in [draft-martini].
- [0211] 6. Upon the completion of the label negotiation, Node-1 and Node-2 will update the data-plane with the label information, that is, to populate the packet filter table 60 and the circuit filter table 80 on the PPE 15.
- [0212] 7. Data flow can now be transmitted over the optical connection.
- [0213] FIG. 18 further explains the process of tearing down or deleting a pseudo wire according to the invention. Essentially, FIG. 18 is a sequence diagram that performs the following processes.
- [0214] 1. The Network Operator sends the deletion requests to both Node-1 and Node-2.
- [0215] 2. Node-1 and Node-2 will exchange control messages and withdraw the labels that are previously allocated for the data flow. In case of SONET connection failure or operational teardown, LDP is responsible for withdrawing labels at provider edges
- [0216] 3. Upon the completion of the operation, Node-1 and Node-2 will update the data plane by deleting the corresponding entries from the Packet/Circuit Filter Tables.
- [0217] FIG. 19 further explains the process of handing outages on the optical connections that affect one or more pseudo wires according to the invention. Essentially, FIG. 19 is a sequence diagram that performs the following processes.
- [0218] 1. The optical connection between Node-1 and Node-2 is no longer working. This could be the result of a planned outage by the carriers, or a link failure in the network. The outage may be detected in any number of conventional fashions and such detection is outside the scope of this invention.
- [0219] 2. Node-1 and Node-2 will update the data-plane immediately. One action is to suspend all the relevant Packet/Circuit Filters on PPE. Another option is to reroute the traffic to another optical connection. The mechanism of rerouting at pseudo-wire level is beyond the scope of this invention.
- [0220] 3. Node-1 and Node-2 will notify the condition to Network Operator.
- [0221] Alternative Architectures Benefiting from Invention
- [0222] The switch fabric 32 is a generalized interconnect between line modules. The interconnects are for optical connections and may also include an additional packet flow interconnect to exchange packet data between modules prior to the mapping engine function. The implementation of the

fabric interconnects is outside the scope the invention and does not impact the invention functions. Conceptually, it is convenient to consider two independent switch fabrics as shown in FIGS. 8 through 13b; the TDM switch fabric 30 for optical connections and the packet fabric 34 for packet data that has not been mapped to an optical connection. However, in practice the interconnect function may be implemented in any fashion and with any number of technologies. Examples of other fabric implementations include a single TDM switch fabric, a single packet switch fabric, and technologies may include any pure electrical, or a hybrid optical/electrical switch fabric.

[0223] Some higher-level architectural details and alternatives will be explored in this section. All of these architectures clearly benefit by utilizing the inventive concepts as further explained below.

[0224] The invention described herein may be implemented on any form of optical connection switch. Given the variety of sizes and designs of switches and the varying needs in data packet capacity requirements, it is natural that there are many possible configurations for incorporating the functionality described in the invention into such switch designs.

[0225] Generally speaking, the functional elements of the switch described herein are not required to be oriented or arranged as shown in FIG. 8. For example, the PPE 15 may be located on a dedicated field replaceable card independent of the line modules 40, switch controller 20, or switch fabric 32 as shown in FIG. 9.

[0226] As further shown in the packet-data-enabled optical connection switch 5' configuration of FIG. 9, the packet server 90 contains the PPE 15 and mapping engine 17 while the MAC 12 is contained on a simplified Packet Access Line Module (PALM'10'). The TDM Line Module 40 is a conventional optical connection originating/terminating module as in FIG. 8. The switch 5' shown in FIG. 9 is a simplified diagram of a practical switch and has only one PALM'10', one TDM line module 40 and one packet server 90 but it is to be understood that in a practical implementation that a plurality of these elements are included to provide the switch with greater capacity.

[0227] Comparing the FIG. 9 configuration of the switch 5' against the FIG. 8 configuration, the mapping engines 17 function identically. The PPE 15 functions are also identical but implementation would be different, thus PPE is labeled 15' in FIG. 9. The switch controller 20 and the tables 25, 60, 80 would also be the same other than differences in switch control coordination of flows to PPE and PPE to optical connection which is more complicated.

[0228] More specifically, the PPE 15' in FIG. 9 sends and receives traffic via the mapping engine 17 and packet fabric 34 while the PPE 15 in FIG. 8 may also send and receive traffic via a physical client port via the MAC 12. In both configurations the PPE's primary function is to manage pseudo-wires in optical connections and translate and manage packet data flows from/to the pseudo-wires.

[0229] In order to benefit from statistical multiplexing gain, many pseudo-wires (on the order of 1,000 s or 10,000 s) will be carried in each optical connection. The data flows that are translated into these pseudo-wires will normally connect to the packet-data-enabled optical connection

switch over many different physical ports. These physical ports may be located on several different PALMs 10. The PPE 15 will aggregate these multiple pseudo-wires and use traffic shaping principals to share one or more optical connections between the pseudo-wires. The source/destination flow associated with each pseudo-wire may reach the PPE via a MAC 12 located on the PALM 10 with the PPE 15, or it may be forwarded via the packet fabric 34 from a PPE 15 located on another PALM 10. This is the architecture shown in FIG. 8.

[0230] As the space and power limits of the PALM 10 will limit the size and capacity of the PPE 15 that can be located on the PALM 10, it may be desirable to locate the PALM on a dedicated module like the packet server 90 shown in FIG. 9. In this configuration, the PPE 15' operates as described above.

[0231] The packet server 90 is essentially another example of switch architecture with the PPE and other data functions included.

[0232] As described earlier, the implementation of the interconnect switch fabric 32 is beyond the scope of the invention. Depending on the implementation of the packet data interconnect function 34, it may be necessary to translate the packet data traffic from/to the PPE 15, 15' into a compatible format for the interconnect. In FIG. 9, the packet fabric interface 16 is fulfilling this function. This is a detail that could be considered part of the packet fabric/interconnect implementation and removed from the figure as in FIGS. 10-13a.

[0233] More specifically, the switch 5' may contain multiple packet server modules 55 to increase the packet processing capacity of the switch 5' and/or for redundancy as shown in FIG. 10. As shown therein, n PALM' modules labeled 10'-1 through 10'-n are provided. In addition, j packet servers labeled 90-1 through 90-j are also provided.

[0234] Packet traffic transmitted between PALM 10' cards and packet server 90 cards can be carried over a packet switch fabric 34 or interconnect as shown in FIG. 10. The packet switch fabric 34 or interconnect may be implemented any number of ways. Examples of implementations include but are not limited to a dedicated packet switch element contained on a field replaceable switching card; dedicated backplane traces between PALMs 10' and packet servers 90; an asynchronous crossbar switch; or dedicated connections between PALMs 10' and packet servers 90 in the TDM switch fabric 30.

[0235] A packet switch fabric 34 or interconnect may be used in the packet-data-enabled optical connection switch 5' even if the architecture does not include packet server modules 90. As shown in FIG. 8, a packet switch fabric 34 or interconnect can be used to transmit packets between PPEs located on multiple PALMs (e.g. between PPE 15-1 on PALM 10-1 and PPE 15-n on PALM 10-n. Transmitting packets between PPEs 15 in such a fashion allows aggregation of packet data from multiple physical ports that reside on different PALMs

[0236] An advantage of a packet-data-enabled optical connection switch 5, 5' is that the same network element can be used to switch a variety of types of traffic. Traditional TDM circuit traffic is switched similarly as on traditional optical connection switches via a TDM fabric such as TDM fabric

32 and TDM line modules 40, 41 as shown in FIG. 12. Simultaneously, the packet-data-enabled optical connection switch 5, 5' can be switching L2 packet flows into pseudo-wires over optical connection as described in the invention and shown in FIG. 13 for the case of a packet server architecture. The PPE 15' and PALM 10' may be implemented to also allow packet switching between packet data ports as shown in FIG. 11.

[0237] As mentioned earlier, an intermediate provider node may have the capability to overwrite an encapsulation label. Such a node would most likely contain a PPE 15 or 15' and mapping engine 17 to perform this function. One reason to overwrite the encapsulation label at an intermediate node would be to aggregate multiple pseudo-wires arriving at the node on different optical connections onto a common outbound optical connection.

[0238] An example of the data path through packet-data-enabled optical connection switch with packet server architecture is shown in FIG. 13a. In this example, an optical connection containing packet data traffic arrives at the switch on TDM line module 40-1 and is switched via the TDM switch fabric 32 to the mapping engine 17-1 located on packet server 90-1. The PPE 15-1 will process the recovered packet as described earlier but the outgoing data interface entry in the circuit filter table will contain a value that reserved for the PPE to loop the packet back into the PPE 15-1 similar to if it were to have arrived from the packet fabric/interconnect 34. The PPE 15-1 will then process the packet again and based on the packet filter table 60 send the packet to the mapping engine 17-1 to go out another optical connection. This other optical connection, originating from the mapping engine 17-1 is switched via the TDM switch fabric 32 to the associated outbound TDM LM, 40-m in FIG. 13a.

[0239] As noted previously, the different types of L2 traffic supported by the packet-data-enabled optical connection switch may require multiple MACs 12 and/or multiple types of PALMs 10, 10'. Additionally, the PALM 10, 10' may contain multiple physical ports that may or may not be sending/receiving the same type of L2 traffic.

[0240] In a general case, a sub-set of ports on the PALM may send/receive conventional TDM optical connection traffic so that the PALM also functions as a TDM LM on a sub-set or all of the traffic. Similarly, a mixture of conventional TDM traffic and L2 traffic may arrive on the same physical port of a PALM. In this case, the L2 traffic is contained in a TDM transport frame that is multiplexed with other transport frames into a single high-speed TDM frame. In order to access the L2 traffic, the PALM 10, 10' would perform conventional TDM add/drop multiplexing (ADM) functionality to terminate the TDM connection containing the L2 traffic and pass the remaining TDM connections to the TDM switch fabric.

[0241] For example, a physical port on a PALM may be receiving/transmitting a SONET OC48 signal with the first 12 STSs carrying ATM traffic and the remaining 36 STSs carrying TDM circuit traffic that is to be switched to other TDM outbound ports on the switch. The PALM 10, 10' would first demultiplex the OC48 signal using conventional means. The resultant tributary that contained the ATM traffic would be terminated and the L2 packets recovered and forwarded to the PPE. The remaining TDM tributaries

would be forwarded to the TDM switch fabric **32**, similar to how they would have been handled had they arrived at the switch on a TDM LM port.

**[0242]** Example Of Inventive Operation

**[0243]** In this section, we walk through an example of how a carrier provisions a pseudo-wire between SONET switches, such as a CoreDirector® (CD) switch made by CIENA Corporation.

**[0244]** As shown in **FIG. 20**, CD-1 (IP loopback address 1.1.1.1) and CD-2 (IP loopback 2.2.2.2) are provided in a network having other (unlabelled) CDs that serve as intermediate nodes in the provider network. A customer attaches to port 1 on CD-1 using VLAN ID **100**, and to port 2 on CD-2 using VALN ID **200**. Inside the SONET transport network, SNC-12 is established ahead of time. SNC-12 can be used to carry Ethernet traffic between CD-1 and CD-2.

**[0245]** Both CD-1 and CD-2 use LDP to discover each other. This allows both nodes to exchange control information to setup the pseudo wires. All control messages are tunneled through SNC-12 as SONET payload and encapsulated with a MPLS “IP4 Explicit NULL Label”.

**[0246]** Once a SNC is in place, establishing a pseudo wire includes three basic steps:

**[0247]** 1. Network Operator Provisioning:

**[0248]** Each VCID uniquely identifies a pseudo wire between a pair of edge nodes. At each node we associate a port/VLAN with a remote edge (loopback address) and VCID. In the example, the network operator picks VCID **50** to identify the pseudo wire between (Port 1, VLAN 100) on CD-1 to (Port 2, VLAN 200) on CD-2. All necessary information is downloaded to CD-1 and CD-2.

**[0249]** 2. MPLS Label Advertisement and Solicitation:

**[0250]** Upon the completion of the provisioning process, LDP automatically exchanges pseudo wire information between CD-1 and CD-2. CD-1 advertises MPLS label 1000 for VCID **50** to CD-2. Similarly, CD-2 advertises label 2000 for VCID **50** to CD-1.

**[0251]** 3. Data Plane Setup:

**[0252]** After MPLS labels have been exchanged, the edge nodes program the data plane for pseudo-wire operation. CD-1 will program the PPE as follows:

**[0253]** For all Ethernet frames received from Port 1 with VLAN 100, push label 2000, and send the frames through SNC-12.

**[0254]** For all Ethernet frames carried over SONET arriving on SNC-12 with label 1000, rewrite VLAN-ID to 100, send them through Port 1.

**[0255]** Similar rules are configured on CD-2 for frames going to CD-1.

**[0256]** Advantages Of Invention:

**[0257]** Martini’s pseudo-wire approach provides a uniformed method to carry all types of layer-2 traffic over a carrier’s backbone network. However, the

backbone must be MPLS/IP-enabled. Traditionally, carriers are very careful with setting up SONET cross-connections inside their networks. In many cases, SONET connections are well provisioned with a rich set of features for network resource allocation, traffic restoration, and link protection, etc. Thus, instead of building pseudo-wires over a MPLS backbone, it would be desirable to use SONET cross-connections to carry pseudo-wire traffic directly.

**[0258]** If backbone networks deliver only layer-2 frames between edges, it may be more economical from both an equipment and management expense point of view to provide the “tunneling” functionality on top of the SONET cross-connections directly, rather than building another layer of tunneling mechanism running on top of optical transport networks.

**[0259]** In the invention, optical transport networks can be used to support both traditional voice traffic as well as data packets. The transport backbones can be provisioned and administrated as they have been for years. Only at network edges, pseudo wires are established to transfer data traffic. Thus, the overall transport management system is not disturbed.

**[0260]** By creating pseudo wires on top of SONET cross-connections, carriers can better utilize network resource by mapping individual user traffic onto SONET virtual concatenated trunks, and adapt mechanisms such as LCAS to fine-tune bandwidth reservations. Since the pseudo wires and the optical cross-connections are originated from the same edge nodes, this can potentially reduce network operation cost for carriers.

**[0261]** The carriers can aggregate data traffic into transport networks directly from network edge. There is no need to introduce UNI or NNI interfaces to bring data traffic into the optical domain. Mapping pseudo-wires into pre-established SNC’s automatically can eliminate the undesired effect of creating and deleting SNC’s dynamically at user and network interfaces.

**[0262]** From a hardware support point of view, this approach will leverage the scalable SONET switching capability in some of the SONET switches. Carriers can bundle and aggregate pseudo-wires into fine-granular STS trunks. It is important to realize that SONET STS trunks themselves are perfect for user flow isolation and bandwidth guarantees. Providing class-of-service or QoS at an STS granularity is hence a unique feature that routers cannot cheaply replace in the foreseeable future.

**[0263]** This invention can aggregate both traditional Layer-2 as well as MPLS labeled traffic over optical transport networks. As a result, this invention can further help network providers to integration services, such as L2 and L2 VPN, more economically.

[0264] Second Embodiment (Admission Control Apparatus, System and Method)

[0265] Terminology of Second Embodiment

[0266] Due to the possibility of a common environment, hardware and application, the second embodiment may use many of the same devices, processes and techniques of the first embodiment. However, it is to be noted that the second embodiment may be applied within a much broader context than the first embodiment. Specifically, the second embodiment may be applied to electrical transport networks that utilize routers and/or L2 switches. Some of the differences are pointed out below in this terminology section while specific hardware and operational differences are explained in following sections.

[0267] CE: Customer Edge. This is also referred as a customer data node or customer edge node throughout this specification. The CE may be a router or a switch.

[0268] PE: Provider Edge. This is a device (also referred to herein as a provider edge node) that either routes or switches traffic. In the context of the second embodiment of the invention, a PE can be a router, a Layer-2 switch or an optical switch.

[0269] Ingress: This refers to traffic entering a provider's backbone from a CE.

[0270] Egress: This refers to traffic leaving a provider's backbone toward a CE.

[0271] Data Tunnel: A data forwarding connection between two PE's. A data tunnel can be a MPLS label-switched-path (LSP), a SONET/SDH cross-connect, or an optical connection. Throughout this second embodiment one of the primary considerations is traffic engineering that may be applied within the backbone portion of the network to data connections to/from the backbone.

[0272] Data Interface: An interface that points to a CE, and is responsible for receiving packets from customer networks.

[0273] Data Flow: A stream of packets that can be uniquely identified through packet headers or the received physical interface. In the context of the invention, each Data Flow will be encapsulated with labels and aggregated into a Data Tunnel. That is, each Data Tunnel can aggregate multiple Data Flows.

[0274] Pseudo-wires: One pseudo-wire maps to one Data Flow. A Pseudo-wire is the Data Flow with all packets encapsulated with a label. This description interchanges the terms Pseudo-wire or Data Flow during for simplicity and clarity.

[0275] Initiating Point: the PE node that initiates the creation of a pseudo-wire.

[0276] Terminating Point: the PE node that terminates a pseudo-wire. Note that each pseudo-wire may be a bidirectional data flow, thus, there may be no significant functional difference between an Initiating and a Terminating point.

[0277] AC Logic: Pseudo-wire Admission Control Logic (see FIG. 24). This logical entity controls the admission control procedure during pseudo-wire setup and tear down. An actual implementation of AC Logic 220 is a software process running on a control module of a router or switch but this AC logic 220 could also be implemented with an ASIC, FPGA, etc as is known in the art.

[0278] The first embodiment described above does not fully address the issues concerning data flow aggregation and the resulting potential for traffic congestion. In the context of pseudo wires used by the first embodiment and also described in conventional pseudo wire techniques such as Pseudo Wire Emulation Edge-to-Edge (PWE3), each data tunnel (such as an optical connection or a MPLS label-switched-path) between two provider edge nodes is capable of aggregating multiple data flows. The aggregation of such data flows can cause real and difficult congestion problems that need to be resolved.

[0279] This congestion problem is magnified for those data flows that require service guarantees from provider backbone networks. If such quality of service guarantees are made then the provider edges should apply some type of admission control to regulate both incoming and outgoing data traffic. Otherwise, the quality of service guarantee cannot be consistently met with the result being that the provider may lose customers, be forced to pay fines, etc. While the need for admission control on incoming traffic flows is apparent and a relatively easy problem to solve, such admission control requirements on outgoing traffic flows can be subtle and tricky to resolve.

[0280] FIG. 23a shows a scenario that exemplifies the need for admission control on outgoing data flows. As shown therein, both Customer Nodes 1 and 3 communicate with Customer Node 2. There exists one pseudo-wire, PW12, between Provider Edge Nodes 1 and 2 to transfer data traffic between Customer Node 1 and 2. Similarly, a pseudo-wire, PW32, is used to carry traffic between Customer Node 3 and 2. PW12 and PW32 require network bandwidth (BW) resources BW\_12 and BW\_32, respectively.

[0281] Within the provider backbone, the provider may deploy techniques such as MPLS in a router backbone, and GMPLS or OSRP (optical signal routing protocol developed by CIENA Corporation) in an optical network backbone to manage the data connections between provider edge nodes. As a result, the packets are not likely to experience any traffic disturbance inside the backbone but this is not true of the ingress and egress data interfaces.

[0282] In the example illustrated in FIG. 23a, both pseudo-wires (PW12 and PW32) exit the network at the same data interface on Provider Edge Node 2. The data interface, therefore, must have enough capacity to handle data traffic that is the "sum" of both pseudo-wires (BW\_12+BW\_32). Otherwise, data traffic from Customer Node 1 and 3 may experience congestion at the data interface. This type of data service is unacceptable, and can be quite costly, particularly considering that data traffic has been well provisioned and delivered within the backbone, but dropped at the last leg of the transmission—the egress interface.

[0283] Note that network resources are generally multi-dimensional vectors and may contain information such as

bandwidth, priority, and service classes. Therefore, it may not be possible to strictly add two resources. For example, one pseudo-wire may call for a higher bandwidth and another calls for a higher priority. In such a case, instead of simply “adding” up two resource vectors, the resource merging routines at the egress interface of the provider edge node must be able to return a third resource vector that is at least as large as each; mathematically, this is the “least upper bound” (LUB). When the term “adding” resources is used herein, it is meant to refer to a LUB operation or equivalent.

[0284] To provide edge-to-edge service guarantees, it is critical to provision network resource on egress provider edge nodes. There are a number of conventional methods to achieve this goal each of which has associated disadvantages that the present invention seeks to avoid:

[0285] 1. Best-effort: This has been assumed in the existing IETF PWE3 framework. This is a reasonable solution if the data service is to deliver best-effort packets only, such as today’s Internet IP traffic. However, this is not acceptable for transporting delay-sensitive traffic, such as voice.

[0286] 2. Over-provisioning: This is a method that over-provisions the egress data interfaces on the provider edge nodes to ensure that needed capacity will always be present. This approach is only reasonable if the providers have the control over the links to the customer networks and where excess capacity is available for the over-provisioning. However, this may not be the case in many existing network configurations, particularly, in situations where the provide networks are transport backbones running SONET/SDH optical connections, and the customer networks are IP router networks. The interfaces between transport and data networks are always administrated separately in current networks thereby making such a method unworkable.

[0287] The present invention proposes a method of exchanging data service information between provider edge nodes that is based upon but which significantly extends the existing PWE3 framework. In the example, at pseudo-wire provisioning time, Provider Edge node 2 will be aware of the resources required for PW12 and PW32, and allocate appropriate resource vectors (including bandwidth capacity) accordingly. If not enough resources are available, the Provider Edge nodes of the invention may apply mechanism, such as preemption and resource shuffling, to make room for more important pseudo-wires.

[0288] Control-Plane Service Negotiation Overview

[0289] FIG. 23b illustrates the operation of service negotiation between Provider Edge Nodes. Provider Edge Nodes C and H aggregate data flows (solid arrow) from Customer Data Node A, B, I and J over a data tunnel (heavy black line). The data tunnel traverses through a number of backbone nodes, D, E and G of the provider backbone network.

[0290] Provider Edge Nodes C and H use signaling protocols, such as those described above in relation to embodiment 1 or by using LDP and draft-martini, to provision pseudo-wires. The result of the provisioning is to aggregate multiple data flows into a single data tunnel. Each data flow is represented as a pseudo-wire within the data tunnel as will be described in more detail below in relation to FIG. 31.

[0291] In addition to the pseudo-wire information that has been described above in the first embodiment, the second embodiment requires the provider edge nodes to exchange the following information for each data flow (or pseudo-wire):

[0292] CIR (Committed Information Rate): This information describes the amount of bandwidth that is required for a given pseudo-wire. This CIR information is essentially the same, in and of itself, as the CIR traditionally used in Frame Relay service offerings but is utilized by the invention in a completely different context and environment. CIR is derived from data flow’s average and peak bandwidth requirements. Each provider may have different CIR settings the description of which is beyond the scope of this invention.

[0293] Class: This information describes the traffic class to which a given pseudo-wire belongs. Although the description of the class may be the same as what has been defined in IETF DiffServ (specifically, RFC2597: Assured Forwarding PHB Group, <http://www.ietf.org/rfc/rfc2597.txt>, RFC2598: An Expedited Forwarding PHB, <http://www.ietf.org/rfc/rfc2598.txt>) and the class format may be the same as DSCP [RFC2474: Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers <http://www.ietf.org/rfc/rfc2474.txt>], the application thereof to the specific provider edge node admission control techniques of the invention is unique.

[0294] Setup Priority: This information describes the priority of a given pseudo-wire with respect to taking resources. This value is used in deciding whether this pseudo-wire can preempt another pseudo-wire. The Setup Priority is assigned by the provider to a customer’s data flow. During pseudo-wire provisioning, when this is no sufficient amount of network resource, a data flow with higher Setup Priority value can preempt the pseudo-wires with lower priority from a data tunnel. The concept is elaborated upon in detail in the operational section below.

[0295] Holding Priority: This invention describes the priority of the pseudo-wire with respect to holding resources. The Holding Priority is used in deciding whether this pseudo-wire can be preempted by another pseudo-wire. The Holding Priority is assigned by the provider to a customer’s data flow.

[0296] Note that the usage of above parameters may exist in other technologies. Specifically, CIR is a concept from Frame Relay; Traffic Class is the central concept in Internet DiffServ service; Setup and Holding Priority have been mentioned in RSVP-TE for MPLS [RFC3209]. However, they have never been bundled together in the context of supporting CE-PE network edge admission control at PE nodes nor has this information been exchanged among PE’s, particularly for the purpose of admission control.

[0297] System Logical View

[0298] FIG. 24 is a high-level block diagram of a packet access line module 210 and control module 200 for performing admission control according to the invention. The elements of FIG. 24 may be included within the provider

edge nodes illustrated in **FIGS. 23a** and **23b**. In this second embodiment, the provider edge nodes may be a router, a Layer-2 data switch, or an optical switch that is capable of processing data traffic. In the latter case of an optical switch capable of processing data traffic, a prime example is the first embodiment of the invention that is described above in relation to **FIGS. 1-22**.

**[0299]** As shown in **FIG. 24**, the inventive system includes a control module **200** that is responsible for handling all control protocol messages including the setup of pseudo-wires. As further shown in **FIG. 24**, the control module **200** includes a pseudo-wire manager **222** and a pseudo-wire admission control logic **220** that interfaces with multiple databases during its operation, which will be explained in detail below. The control module **200**, pseudo-wire admission control logic **220** and pseudo-wire manager **222** may all be implemented as software processes and executed by an available microprocessor which may already be present in the provider edge node. Alternatively, these elements may be part executed by a separate microprocessor or group of microprocessors or may be implemented by an FPGA, ASIC or other combination of hardware and software as is generally known in the art.

**[0300]** As further shown in **FIG. 24**, the invention further includes a packet-access line module **210**. Multiple packet access line modules **210** are typically included in an actual implementation but for the ease of explanation only one is illustrated in **FIG. 24**.

**[0301]** The packet access line modules **210** differ from the PALMs **10, 10'** of the first embodiment in that the packet access line modules **210** are more general and are not so tied to the SONET/SDH optical switching environment of the first embodiment. In the first embodiment, the PALM **10, 10'** is responsible for processing MAC packets and maps them into an optical connection, thus, it is a component on an optical switching device. Whereas in the second embodiment, a PALM **210** can process packets from any data interface, and send them out to any data interface. Thus, the PALM **210** of the second embodiment can be one used on routers. However, in both cases, the PALM needs to be capable of supporting packet admission control, such as policing and shaping.

**[0302]** The packet access line modules **210** interface with Customer Equipment (CE) Data Nodes, and are responsible for aggregating data flows using a multiplexer (not shown) or other conventional device for the physical aggregation. This could be a line module that interfaces with Customer Equipment Data Nodes directly, or a service module that process data flows from other line modules within the same system. Nevertheless, each Packet-Access Line Module **210** processes packets coming from multiple data interfaces, and aggregates them into multiple outgoing interfaces toward the provider's backbone, which is explained below in more detail.

**[0303]** To interface with the provider backbone, one or multiple provider-interface line modules (not shown) should be provided as is known in the art. Such provider-interface line modules interface with the Packet-Access Line Modules **210** to inject packets into provider's backbone in the conventional fashion. Depending on the type of provider's network, they can be a conventional router packet forwarding module, a Layer-2 switching module, or a TDM switching module (see first embodiment above for an example).

**[0304]** Using or otherwise accessing the control module **200**, network operators request the setup of pseudo-wires. The actual details of pseudo-wire setup have been exemplified and specified in other relevant documents [draft-martini, LDP] a significant variation of which is described above in relation to embodiment one. To support the second embodiment, the pseudo-wire process needs to maintain the following additional information within the databases shown in **FIG. 24**:

**[0305]** Session Table **225**: The session table **225** may be stored in a control message database **218** to maintain all the logical connections that the system has currently established with other provider edge (PE) nodes. Note that each PE node may have multiple parallel connections to another PE node, which results in having multiple logical peering sessions. Each peering session may be a TCP session, which can be uniquely identified by the combination of IP source and destination address, and source and destination port numbers. Further details of the session table **225** are discussed below in relation to **FIG. 28**.

**[0306]** Mapping Tables: A provider-edge node system of the invention maintains two mapping tables within a mapping database **219**: a packet filter table **260** and a circuit filter table **280** are maintained by the control module **200** to enable admission control. The packet filter table **260** is maintained to regulate data flows coming into the backbone (from a customer edge node), while the circuit filter table **280** is for controlling data flows leaving the backbone (and to be provided to a customer edge node).

**[0307]** Resource Tables: The control module **200** also includes a pair of resource tables including an ingress resource table **232** stored in an ingress database **230** and an egress resource table stored in an egress database **235**. These resource tables respectively maintain all data tunnels (provider-bound) and data interfaces (CE-bound) on a PE node. A data tunnel can be a MPLS Label-Switched-Path, a SONET/SDH cross-connection, or an optical DWDM connection. A data interface can be any interface that can carry data packets, such as Ethernet, FDDI, ATM, POS etc. In the context of the invention, all the data tunnels and data interfaces described here are always associated with network resources such as link bandwidth. There is much less relevance for using admission control in networks and links that do not support traffic engineering or QoS (quality of service).

**[0308]** The pseudo-wire controller **215** on each packet access line module **210** is responsible for processing the actual packets. System-wide control is provided by the pseudo-wire manager **222** that keeps track of system-wide information (e.g. the control message database **218** and session table **225** have system-wide information for the pseudo-wire manager **222**). In contrast, the pseudo-wire controller **215** within each of the Packet Access Line Modules **210** maintains subsets of Mapping Tables and Sessions Tables.

**[0309]** The packet access line modules **210** also include an ingress process **212** that interfaces with an appropriate

subset (260-1) of the packet filter table **260** to control incoming data flows for that packet access line module **210**. Similarly, an egress process **217** interfaces with an appropriate subset (**280-1**) of the circuit filter table **280** to control outgoing data packets for that packet access line module **210**.

[0310] To support pseudo-wires, the ingress process **212** is responsible for encapsulating labels to incoming packets, that have a match in the Packet Filter Table **260**, while the egress process **217** strips off the labels when delivering those packets to the customer networks. However, both processes may share the same set of traffic conditioning mechanisms in regulating incoming traffic. Some of the typical and conventional traffic conditioning mechanisms that may be used by the invention include RED (“Random early detection gateways for congestion avoidance”, Sally Floyd, Van Jacobson, IEEE Transaction Networking, 1993), Token Bucket, protective buffer management (“Protective buffer management policies”, I. Cidon, R. Guerin, and A. Khamisy, IEEE Trans. Networking, 1994), and WFQ (“Analysis and simulation of a fair queuing algorithm”, A. Demers, S. Keshav and S. Shenker, Journal of Internetworking, 1990). By applying these mechanisms, packet traffic flows will behave according to the QoS parameters defined between PE’s.

[0311] The pseudo-wire controller **215** generally operates as follows: When sending control messages within data tunnels, pseudo-wire controller **215** relies on the session table **225** to determine where to forward the control messages. In other words, the pseudo-wire controller **215** generally operates like the PPE controller **70** of the first embodiment details of which are provided above.

[0312] Before turning to details of the inventive operation, the tables and databases used by the invention will be discussed.

[0313] Packet Filter Table **260**

[0314] The packet filter tables are for the purpose of handling packets coming from the customer equipment (CE) on the ingress interfaces. FIG. 26 illustrates a packet filter table **260** that is relevant to and used by the invention. Like the packet filter table **60** of the first embodiment, the packet filter table **260** of the second embodiment includes several data fields in common, namely, the data interface, label, data tunnel, and encapsulation label data that are described above in more detail. The packet filter table **260** of the second embodiment, however, includes additional data fields as further described below:

[0315] A Searching Key which includes the packet’s (incoming) data interface and label information.

[0316] (Incoming) data interface: This is the interface that receives the packet. It can be the identification for either a physical or logical interface. The invention makes no assumption on how such information is actually obtained. However, the interface information is required for each packet being received. In some applications, the system can direct incoming traffic base on the incoming data interface information only.

[0317] Label: This can be, for example, a Layer-2 header. A Layer-2 header can be an Ethernet MAC and VLAN tag, a Frame Relay DLCI, or an ATM VCI/VPI.

[0318] Data Tunnel: This is the connection that the packet will be injected into as it enters the provider network. As shown in the figure, a data tunnel can be a MPLS LSP, a SONET Virtual Concatenation path, or an optical cross-connect.

[0319] Encapsulation Label: The label for each data flow. It will be encapsulated with the packet. The encapsulation label can be in MPLS label format, or any other format, so long as it can uniquely identify each individual data flow (that is, pseudo-wires) within a data tunnel.

[0320] CIR (Committed Information Rate): This indicates the amount of network bandwidth that an incoming data flow can consume. If incoming traffic exceeds this value, traffic congestion may result within a data tunnel. For packets that do not comply, a user-defined traffic conditioning mechanism, such as RED, WFQ etc as described above and buffer management, will be used.

[0321] Class: This indicates the traffic class that incoming packets belong. The traffic classes can be Assured Forwarding (AF) and Expedited Forwarding (EF) classes defined in RFC2597 and RFC2598 referenced above in which, there are four independent AF classes, and one EF class.

[0322] Setup Priority: The priority of the pseudo-wire with respect to taking resources. This value is used in deciding whether this pseudo-wire can preempt another pseudo-wire.

[0323] Holding Priority: The priority of the pseudo-wire with respect to holding resources. The Holding Priority is used in deciding whether this pseudo-wire can be preempted by another pseudo-wire.

[0324] The CIR, Class, Setup and Holding Priorities of the packet filter table **260** may be assigned by the provider. As shown in FIG. 26, the packets from Filter-1 and Filter-3 share the same data tunnel, SONET VCG Number-3. Within the data tunnel, each data flow is differentiated by the encapsulation label. The packets that belong to Filter-1 and Filter-3 belong to the same traffic class, AF-1, however, each has a different CIR value. To avoid potential traffic congestion, the PE will apply a conventional mechanism such as RED and buffer management to regulate ingress traffic flows.

[0325] Circuit Filter Table **280**

[0326] FIG. 27 provides further details of the circuit filter table **280** that is used by the second embodiment to handle packets coming from the provider backbone and going toward CE’s at the egress interfaces. As shown there, circuit filter table **280** includes the following attributes many of which are shared with the circuit filter table **80** of the first embodiment:

[0327] A Searching Key which includes the provider backbone-bound data tunnel and label information.

[0328] Data Tunnel: The data connection where the packet arrives from the backbone. It can be a MPLS LSP, a SONET VCG (Virtual Concatenation Group), or an optical interface.

- [0329] Label: This is the label that has been inserted at the ingress of the data flow. It is used to identify a specific data flow within a data tunnel.
- [0330] Outgoing Data Interface: The interface where the packet is to be forwarded.
- [0331] CIR (Committed Information Rate): This indicates the amount of network bandwidth that an outgoing data flow can consume. If data traffic exceeds this value, traffic congestion may result at an egress data interface. For packets that do not comply, a user-defined traffic conditioning mechanism, such as RED, WFQ, and buffer management, will be used.
- [0332] Class: This indicates the traffic class that outgoing packets belong. The traffic classes can be Assured Forwarding (AF) and Expedited Forwarding (EF) classes defined in RFC2597 and RFC2598, in which, there are four independent AF classes, and one EF class.
- [0333] Setup Priority: The priority of the data flow with respect to taking resources. This value is used in deciding whether this data flow can preempt another one.
- [0334] Holding Priority: The priority of the data flow with respect to holding resources. The Holding Priority is used in deciding whether this data flow can be preempted by another one.
- [0335] In the circuit filter table 280 example of FIG. 27, Filter-1 and Filter-4 share the same outgoing data interface (data port 1). Packets from each filter must comply with the CIR that has been negotiated during the setup of the pseudo-wires.
- [0336] Session Table 225
- [0337] The session table 225 is used to keep track of all the control information with peering PE's. In the context of the invention, the provider can assign network resources to ensure the reliable and timely delivery of the control messages.
- [0338] The session table 225 applies to control messages that are delivered as special "labeled" packets within the data tunnels where pseudo-wires will traverse details of which are described above in relation to the first embodiment.
- [0339] As shown in FIG. 28, session table 225 has the following attributes many of which are shared with the session table 25 of the first embodiment:
- [0340] Searching Key: Control Message ID
- [0341] Each control message carries a unique ID to identify the "peering session" to which it belongs. A "peering session" is a logical connection between two edge nodes and is used to exchange control information between two nodes. For example, in pseudo-wire operation, the customer may apply LDP [RFC3036]-Label Distribution Protocol, <http://www.ietf.org/rfc/rfc3036.txt> to negotiate data flows. LDP operates over TCP. Between two edge nodes, all control messages go over a TCP session that can be uniquely identified with TCP Sender Port Number, and IP addresses. In this invention, the exact message ID format is not specified but is left open to provide compatibility with existing systems. However, it is reasonable to assume that each control message carries enough information to identify the session to which it belongs.
- [0342] As an example, FIG. 28 shows three sessions that are identified with, in this non-limiting example, TCP port numbers.
- [0343] Outgoing Data Tunnel: This is the connection that the control messages will be injected into.
- [0344] Encapsulation Label: The identifiable label for the control message. The system will insert this label to the control message as further described above in relation to the first embodiment.
- [0345] CIR (Committed Information Rate): All control messages within a session will have a fixed network resource level. This is designed to protect the control messages from potential congestion caused by regular data traffic.
- [0346] Class: Control messages are assigned to a traffic class, according to DiffServ. This will further improve the latency for packet delivery. In the example, we assign EF (Expedited Forwarding) class to all three control messaging sessions. Consequently, control messages will always have the highest priority over regular data packets.
- [0347] Note that in FIG. 28, there are three peering sessions on a specific PE node that are differentiated by TCP source port. In an actual implementation, these three peering sessions correspond to three different TCP sockets on a POSIX interface. It is possible that there is no CIR assigned to a control message for a number of reasons:
- [0348] 1. There is plenty of bandwidth within the data tunnel.
- [0349] 2. The system relies on transport layer (TCP retransmission) for reliable message delivery.
- [0350] In the example of FIG. 28, Session-3 has no CIR; however, it does have an EF class to ensure the timely delivery of the control messages. This illustrates some of the variations on how the session table may be structured and utilized by the invention.
- [0351] Resource Tables
- [0352] The control module 200 of each provider edge node maintains local resource usage information concerning both provider-network-bound data tunnels and customer-equipment-bound data interfaces. These inter-related resource tables are shown in FIG. 29a and FIG. 29b and respectively illustrate the ingress resource table 232 and the egress resource table 237 maintained by the control module 200.
- [0353] The resource tables maintain the available resource information for each data flow local to the PE node in which the control module 200 resides. As shown in FIGS. 29a and 29b, the inter-related resource tables 232, 237 support multiple service classes and the control module 200 keeps track of the available resource information per each traffic class.



[0354] Furthermore, the inventive resource tables 232, 237 support multiple types of PE/CE data interfaces including ATM, Ethernet, Frame Relay, PPP, RPR, Ethernet over SONET (EoS), LAPS, GFP-F, and Cisco-HDLC. As such, the inter-related resource tables of the invention may be termed multi-service-class/flexi-interface resource tables that may be used by provider edge nodes to negotiate consistently managed data tunnels across a provider network on behalf of data flowing from/to a diverse base of customer edge nodes.

[0355] The resource tables 232, 237 shown in FIGS. 29a and 29b provide a typical but non-limiting example of the invention and its many applications. Although three service classes are shown the number could vary. Some granularity in the number of service classes is desirable both from a standpoint in traffic negotiation but also in terms of creating different levels of service points necessary for a broad range of network offerings. The invention further improves these offerings by permitting network providers to make better use of their networks by admitting more pseudo-wires and, thereby, more customer traffic. Thus, not only does the invention permit customer traffic from diverse data interfaces to have multiple service classes, the invention also permits the provider network to more efficiently utilize all of the available bandwidth to carry this traffic.

[0356] Further details of the resource tables 232, 237 will become apparent in the following sections.

[0357] Operation Of Second Embodiment

[0358] The following sections describe the methods and operations of the second embodiment. These methods may be performed utilizing the elements described above or their equivalents.

[0359] Provisioning Pseudo Wires for Admission Control

[0360] This section describes the inventive method of provisioning pseudo-wires that permit admission control functionality. Following sections will describe both pseudo-wire shuffling and preemption according to the inventive concepts.

[0361] The provisioning of pseudo wires that will permit admission control functionality to operate is a process that includes the exchanging of specific resource information between PE nodes during the creation of pseudo-wires.

[0362] Setting up pseudo-wires (PW) may follow a procedure as defined in [PWE3-CTRL] or the procedure as defined above in the first embodiment. The second embodiment, however, modifies these provisioning processes to operate in context of and to otherwise permit admission control functions over pseudo-wires and in a more general context that is not necessarily limited to optical transport networks.

[0363] FIG. 25a FIG. 25b are high-level flowcharts illustrating the processes and methods performed by the invention for pseudo-wire admission control provisioning from the perspective of initiating and terminating points. For ease of illustration, the method in both figures operates on two PE nodes, each connected to a CE node. For reference sake, one of the PE edge nodes is called the Initiating PE and the other the Terminating PE.

[0364] As shown in FIG. 25a, the carrier/provider network requests to set up a data flow on the Initiating PE. The

Initiating PE receives the message on its control module 200. The carrier network request includes the network resource information [CIR, Class, Setup Priority, Holding Priority] for the new data flow.

[0365] The AC control logic 220 may then determine (350) if there is enough bandwidth available of the data tunnel to support the new data flow connection request by referring to the ingress resource table 232. This determination (350) also includes the AC control logic 220 searching its packet filter table 260 to determine which data flows already exist that are configured to utilize the corresponding data tunnel. The CIR for all flows is tallied and the switch controller determines how much spare BW there is available on the data tunnel.

[0366] If the new data flow CIR would not exceed the data tunnel's capacity, the pseudo-wire manager 222 will negotiate with the corresponding Terminating PE node to continue the creation of the pseudo-wire (358).

[0367] If the addition of the requested data flow would cause the data tunnel to exceed its resource, then the AC control logic 220 searches the packet filter table 260 to determine (352) if the request can be accommodated by adjusting the existing flows toward the terminating point. Such adjustment may include preemption and shuffling as described in detail below. If the adjust can gather enough resources on the data tunnel for the new data flow, the AC control logic 220 will proceed to adjust (356) the flows. The pseudo-wire manager 222 may then continue the negotiation (358) of pseudo-wire with the terminating PE node. Otherwise, the AC logic 220 will refuse or otherwise deny (354) the connection.

[0368] Upon the completion of pseudo-wire creation, the AC logic 220 updates (360) the packet filter table 260 as well as the ingress resource table 232 with the network resource information for the new data flow.

[0369] At the terminating PE node, as shown in FIG. 25b, the control messages are captured and forwarded (370) to the local control module 200.

[0370] The termination PE node control module 200 will determine (374) the resource of the corresponding data port using the egress resource table 237 and by searching the circuit filter table 280 to determine the total CIRs of all pseudo-wire connections terminating on the given data port. If the requested data flow CIR would not exceed the data port capacity, the AC logic 220 will accept the new pseudo-wire.

[0371] If the addition of the requested data flow would cause the data port to exceed its resources, then the switch controller searches the circuit filter table 280 to determine (376) if the new flow can be accommodated by adjust the existing pseudo-wires, which we will describe in detail below. If more resource can be found, AC logic 220 will proceed with the resource adjustment (380) action. Otherwise, the AC logic 220 will deny (378) the connection.

[0372] Upon accepting a new flow, AC logic 220 will update (384) the corresponding entries in the circuit filter table 280 and egress resource table 237. As a part of pseudo-wire process as defined in LDP and draft-martini, the Terminating PE node will acknowledge (382) the establishment of the new flow to the Initiating PE node.

[0373] Note that FIGS. 26a and 26b are very similar in dealing with the data flow. However, FIG. 26a is the logic to admit flows into data tunnels in the backbone, while FIG. 26b is the logic dealing with the data interfaces going toward the customer networks.

[0374] FIG. 30 is a mid-level flowchart illustrating processes and method that may be performed by the invention for pseudo-wire admission control provisioning at both pseudo-wire initiating and terminating points (the PE nodes) and further including both pseudo-wire shuffling and pseudo-wire preemption processes to maximize the number of pseudo-wires admitted. The separate processes performed are indicated by the dashed line separating the respective initiating (labeled transmitting in the figure) initiating and the terminating (labeled receiving in the figure) PE node processes. Since the communication may be bi-directional the terms transmitting and receiving are not accurate in all circumstances but these terms are helpful to gain an understanding of the invention. The relevant components of the control module 200 and packet access line module 210 in the PE nodes that perform these methods are shown in FIG. 25.

#### [0375] Initiating Point Operation

[0376] At the initiating point, Pseudo-wire Admission Control Logic 220 (AC Logic 220) will initiate (700) a pseudo wire with another PE node. To do this, the AC logic 220 first determines (705) if there remains a sufficient amount of resources between the two PE nodes to satisfy the pseudo wire to be initiated. The determination (705) involves an extensive search in the Ingress Resource Table 232 (see FIG. 29a) for the available bandwidth. As mentioned above, network resource referred to here may be a multi-dimensional vector that includes bandwidth and traffic class therefore the determination is a class-by-class determination for available bandwidth that is facilitated by the ingress resource table 232.

[0377] It is possible and likely that there are multiple parallel data tunnels between two PE's. In this case, the provider may set up a pseudo-wire on any of the data tunnels (links) between the two PE nodes that can satisfy the resource requirement. The actual link selection may depend on the network provider's policy, which is a topic beyond the scope of this invention.

[0378] If there are not enough network resources, the AC Logic 220 will attempt to shuffle (710) pseudo-wires on all the links between two PEs. The goal of the shuffling (710) is to free up link resources to accommodate the new pseudo-wire. The mechanism for shuffling (710) will be described below. It is possible that the provider may not allow the practice of pseudo-wire shuffling to avoid traffic disturbance on operational data flows that may occur in some instances. It is also important to recognize, however, that such a potential traffic disturbance may be avoided by the invention. This may be done by the AC Logic 220 which, in the case that it determines that shuffling pseudo-wires would not free up enough resource for the new flow, would not permit shuffling to take place.

[0379] If it is determined (720) that shuffling (710) does not free up enough network resources, the AC Logic 220 will try to preempt (725) less important flows to make room for the new flow. We will describe the detailed preemption procedure (725) below. Note that the policy that determines

or assigns relative importance levels to a flow is internal to network providers and their policies, and is also beyond the scope of this invention.

[0380] If it is determined (730) by AC Logic 220 that it cannot gather enough network resources by preemption (720), it will reject (735) the user's request for a new pseudo-wire creation. Otherwise, after the preemption (720) of the less important flows, AC Logic will once again shuffle (740) the remaining pseudo-wires and make enough room for the new flow.

[0381] Finally, AC Logic 220 formally admits the new pseudo wire flow by updating (745) the corresponding entries in ingress resource table 232 and packet filter table 260. The system will then continue the creation of the new pseudo-wire by exchanging control messages with the terminating node (750). The control messages may be routed to the terminating point via routers over the backbone network. Alternatively, the control message may be switched by pseudo-wire manager 222 to the provider backbone network via pseudo-wire controller 215 and egress process 217. In the latter case, the reliability and performance of the message delivery could be guaranteed by allocating network resources to control message traffic, as indicated in Session Table (FIG. 28). A particular and preferred example of such control messaging is described above in the first embodiment.

#### [0382] Terminating Point Operation

[0383] The control message from the initiating provider edge node is received (755) by the provider edge node that will serve as the terminating point. The new pseudo wire request is extracted from the encapsulated control message in a manner like that described above in relation to the first embodiment with a unique label identifying the message as a control message such that it may be routed to the pseudo-wire manager 222 and AC logic 220. Upon receiving (755) the new request from the initiating point, the terminating point's AC Logic 220 will determine (760) if the outgoing data interface has available resource to accommodate the new pseudo-wire.

[0384] If there are not enough network resources, the AC Logic 220 will try to preempt (765) the less important flows. If the pre-emption process (765) fails to free up enough network resources for the new pseudo-wire as determined (770) by the AC logic with reference to the egress resource table 237, AC Logic 220 will reject (775) the setup of the pseudo-wire. The rejection process (775) includes sending a control message back to the initiating PE node such that initiating node may update its packet filter table 260 and ingress resource table 232. Otherwise, the AC Logic 220 of the terminating node will admit the new flow by updating (780) its corresponding circuit filter table 280 and egress resource table 237.

[0385] Note that no shuffling or preemption should take place on pseudo-wires, unless AC Logic 220 determines that such action would gather enough resources for the new flow. Likewise, the pseudo-wire managers 222 of the initiating and terminating points should only trigger AC Logic 220 when there is a high probability that the new pseudo-wire would pass other checks, and be created successfully. Otherwise, it will cause unnecessary traffic disturbance to the existing data service.

[0386] Pseudo-Wire Shuffling

[0387] By applying pseudo-wire shuffling and preemption techniques described below in more detail, network providers can make better use of their network resources by admitting more pseudo-wires.

[0388] FIG. 31 diagrammatically illustrates the concept of pseudo-wire shuffling according to the invention.

[0389] In the example shown therein, there are two parallel links (data tunnel 1 and data tunnel 2) between two PE nodes that together support five flows, Flow-1 to Flow-5. Before the new flow (Flow 6) arrives, data tunnel 1 carries Flows 1, 2 and 3. When a new flow, Flow-6, arrives, there are not enough resources on either of the links to admit the new flow. To admit Flow-6, AC Logic 220 can move Flow-3 from Data-Tunnel-1 to Data-Tunnel-2 as further illustrated in FIG. 31. After the moving, Flow-6 may be accommodated with Data-Tunnel-1. The invention refers to this “pseudo-wire shuffling” or “shuffling pseudo-wires.”

[0390] It is to be noted that FIG. 31 is a bit simplistic in that the only network resource being fully illustrated is bandwidth while the invention is capable of handling bandwidth demands within a plurality of service classes. Nevertheless, FIG. 31 is helpful for understanding the inventive concept of shuffling pseudo-wires as that term applies to the full network resource vector.

[0391] The process of shuffling, in and of itself, is conventional. In fact, a similar practice has been done in plain old telephone systems (POTS) for years. However, the present invention represents a new application of the shuffling technique within the context of data (IP) networks in general and pseudo-wires in particular.

[0392] In addition, each shuffling requires extensive PE-node-to-PE-node negotiation on the “shuffled” flow. In FIG. 31, the moving of Flow-3 would trigger a deletion procedure in Data-Tunnel-1, followed by a creation procedure in Data-Tunnel-2. Both procedures require extensive control-plane message negotiation which may introduce impact on existing data flows and the performance of the control plane. There are mechanisms to alleviate the stress imposed by the shuffling one of which is described below but others of which are certainly applicable to the invention.

[0393] FIG. 32 is an algorithm to shuffle pseudo-wires. In other words, the shuffling process of FIG. 32 is an example of how the shuffling steps 710, 740 (FIG. 30) may be performed.

[0394] As illustrated in FIG. 32, based on the new flow’s inputted (800) resource requirement and destination PE node identification, the algorithm will first search (805) the session table 225 (FIG. 28) to check if there exists at least one parallel link (data tunnel) to the destination PE. If it is decided (810) that there is no parallel link (not more than one data tunnel), the process will terminate (815) since there is no place to shuffle old flows into. In other words, the shuffling process has failed and the new flow is not admitted.

[0395] If there is more than one data tunnel that the invention can use to shuffle flows, the algorithm will search (820) the packet filter table (FIG. 26) and Ingress Resource Table 232 (FIG. 29a) to determine if enough network resources can be found after shuffling one or multiple existing flows. The actual searching (820) procedure varies

depending on network utilization, network topology, user traffic behavior, and provider policy. For example, shuffling could start on the data tunnel with the most available resource, or the one with the most physical bandwidth. Another implementation issue would shuffle the flows with the least resources first, vs. the ones with the most resources. Also, the selection of the shuffling flows can be sorted or random.

[0396] If the algorithm determines (825) that there are not enough network resources to accommodate the new flow, the algorithm will terminate (815) the shuffling process. Otherwise, the search procedure (820) will produce a list of pseudo-wires that may be shuffled by the shuffle step 830. Shuffling (830) a flow may follow the following sequence:

[0397] 1. If possible, direct existing traffic to a backup link (e.g. such as using protection bandwidth triggered via a conventional APS (automatic protection switch) protocol for SONT/SDH traffic).

[0398] 2. Negotiating with the destination PE to create a new pseudo-wire over the data tunnel where the flow will move into. The negotiation can be the same pseudo-wire setup procedure described in LDP, draft-martini or FIG. 30 above.

[0399] 3. If possible, upon the completion of the new pseudo-wire, move data traffic into the new data tunnel from the backup link.

[0400] 4. Notify the destination PE to withdraw the pseudo-wire from the old data tunnel. The negotiation can be the same pseudo-wire withdrawal procedure described in LDP, draft-martini or FIG. 30 above.

[0401] 5. Update the corresponding entries in the Packet Filter Table 260 (FIG. 26), and Ingress Resource Table 232 (FIG. 29a).

[0402] Note that in the above sequence, shuffling a flow is preferably a “make-before-break” process, which does not impact the ongoing user traffic.

[0403] FIG. 34 shows the operational sequence of pseudo-wire shuffling between two PE nodes in terms of both the data plane and control plane. It is assumed that there are multiple data tunnels and one pseudo-wire (PW1) between PE1 and PE2 initially. When a new pseudo-wire PW2 is admitted into the network, PE1 needs to shuffle PW1 into another data tunnel by first backing up PW1’s traffic. It follows by setting up a new pseudo-wire on a different data tunnel. After PW1’s traffic is redirected into the new pseudo-wire, PE1 can begin the process of setting up PW2 as further illustrated in FIG. 34.

[0404] Pseudo-Wire Preemption

[0405] Preemption, in and of itself, is a conventional technique for CAC (Call Admission Control). The general idea is to rank the importance, or priority, of a flow relative to the others competing for admission into a network. Priority considerations are utilized when a set of flows attempting admission through a node or a link that cause overbooking of resources. CAC resolves the overbooking or oversubscription problem by rejecting one or more of the flows competing for admission. Network nodes also use priorities to preempt some previously admitted low-priority

flows in order to make room for a newer, higher-priority flow. The application of such CAC techniques in the specific environment disclosed herein, particularly in combination with the other features of the invention is a novel and highly advantageous features.

[0406] In the invention, two basic priority classes may be used for each flow: Setup Priority and Holding Priority which are defined above in detail and further discussed below.

[0407] Setup Priority is the relative importance (ranking) of a new pseudo-wire with respect to taking resources from other pre-established pseudo-wires.

[0408] Holding Priority is the relative importance (ranking) of an existing pseudo-wire with respect to holding the resources from being taken away or pre-empted by another pseudo-wire requesting admission.

[0409] For any data flow at an ingress and egress point, its Setup Priority is preferably less than or equal to the Holding Priority. The gap between Setup and Holding Priority makes it harder for a data flow to preempt others, but once it succeeds, the higher Hold Priority makes it easier for the flow to prevent being preempted itself. This mechanism provides a mechanism for balancing between dependency and priority.

[0410] Both Setup and Holding Priorities may be assigned by the providers or network operators using a craft interface, network administrator node, or other known technique. The preemption algorithm of the invention applies at both ingress and egress interfaces of a pseudo-wire, as illustrated in the flowchart of FIG. 30 (e.g. see pseudo-wire preemption step 725 (transmit or ingress interface) preemption step 770 (receive or egress interface)).

[0411] The preemption steps illustrated in FIG. 30 and discussed above may utilize the more detailed preemption algorithm illustrated in FIG. 33. The preemption algorithm requires an input (900) providing the information on a pseudo-wire's resource requirement, set-up priority and the data tunnel or interface where preemption will take place. Note that this algorithm is applied at both ingress and egress interfaces. At ingress, preemption may take place in a data tunnel, whereas preemption can remove less important flows from a data interface at an egress point.

[0412] The algorithm begins in earnest by searching (905) for all the flows having a holding priority less than the setup priority of the new flow. If at the ingress point, the searching (905) is done within the packet filter table 260.

[0413] With this information in hand, the method may then determine (910) if the combined resources from the selected flow(s) having a holding priority less than the setup priority are not enough to accommodate the new flow. If so, the algorithm will fail (915) the preemption process because the new flow simply cannot be accommodated according to the relative priority levels and resource demands. Otherwise, the algorithm will select (920) a set of flows that can accommodate the new flow and which may be preempted according to the relative setup and holding priorities. The combined resources from the selected (920) flows will be larger or equal to the resources required by the new flow. The

actual selection mechanism (920) may be based on provider policy. For example, only the smallest flows would be preempted, or the flow selection can be random.

[0414] After the selection (920) of the preempting flows, the algorithm will notify (925) the PE nodes corresponding to the flows being preempted in order to trigger the pseudo-wire withdrawal procedure (defined in LDP and draft-martini) that removes the flows from the control-plane. The method should also notify (930) the network operator of the preemption events. After which or at the same time, the algorithm will update the corresponding entries in packet filter table 260, circuit filter table 280, ingress resource table 232 and egress resource table 237 as appropriate and thus remove (935) the flows from the data-plane.

[0415] FIG. 35 and FIG. 36 are the operational sequence diagrams for pseudo-wire preemption at ingress and egress, respectively. These diagrams show the operations relative to both the control plane and data plane relative to the initiating and terminating provide edge nodes 1 and 2. As detailed therein, the creation of a higher priority PW2 will result in the deletion of PW1.

[0416] It is important to realize that both shuffling and preemption can be applied during pseudo-wire modification as well and are not limited to pseudo-wire creation. For example, a user may decide to increase the bandwidth allocated to a pseudo-wire. The AC Logic 220 at both ingress and egress will apply appropriate mechanisms described in this invention to accommodate the modification request in much the same way as the creation or initiation of a pseudo-wire.

[0417] Third Embodiment

[0418] The invention also includes a third embodiment that is generally directed to pseudo-wire probing. The pseudo-wire probing feature allows the network operators to actively detect the aliveness of pseudo-wires from network edge.

[0419] The general concept is derived from LSP-ping [LSP-PING], which supports both RSVP-TE and LDP. All probing packets may use UDP (port 3503).

[0420] The existing LSP-ping is to "ping" at the FEC level. The processing procedure can be very complex due to LSP merging among nodes inside the network. Worse, for load-balanced traffic, the LSP-ping cannot probe the intended path accurately.

[0421] To probe pseudo-wires, the invention modifies the scope of the conventional protocol and simplifies the implementation. The invention essentially probes edge-to-edge, point-to-point connections. For example, instead of "pinging" at FEC level, the invention will "ping" per VCID (that is, per pseudo-wire).

[0422] Since each pseudo-wire is always strictly a point-to-point connection between two network edges, the probing will always be accurate. Thus, the protocol level processing is largely simplified.

[0423] At the data processing level, at ingress, the invention marks the probing messages and injects them into the targeted pseudo-wires. The probing messages must be processed differently at the PPE. Each probe message is encapsulated

sulated with a special control word beneath the MPLS header as described above in relation to the first embodiment.

[0424] The procedure of the third embodiment is as follows:

[0425] The control module will provide the following to the PALM:

[0426] probing message payload;

[0427] message length;

[0428] the targeted pseudo-wire ID.

[0429] On the PALM, the CPU and the PPE will perform the following method:

[0430] 1. Based on the pseudo-wire ID, the CPU constructs a MPLS header and a layer-2 header to the probing message. The MPLS label and the layer-2 id's (such as VLAN-tag, DLCI, etc.) must be the same as the ones used by the pseudo-wire that carries user traffic.

[0431] 2. In addition, a Control Word needs to be created and inserted beneath the MPLS header.

[0432] 3. Update the message length field in the layer-2 header.

[0433] 4. Send out the control messages through the same SNC used by the user traffic.

[0434] At the egress edge, the PPE performs the following method:

[0435] 1. After checking on the incoming MPLS label, check if there exists a control word that indicates the packet is a probing message.

[0436] 2. If the result is negative, forward the packet to the corresponding outgoing port.

[0437] 3. Otherwise, forward the packet to the CPU, which in turn will send it up to the control module.

[0438] It is to be understood that the inventive concepts are not limited to SONET and also include SDH which is the prevailing standard in Europe and emerging standards such as OTN. In other words, although the invention (particularly the first embodiment) is described mainly in terms of SONET in the interest of simplifying the description, the inventive concepts may be equivalently applied to SDH or OTN networks.

[0439] The invention being thus described, it will be obvious that the same may be varied in many ways. Such variations are not to be regarded as departure from the spirit and scope of the invention, and all such modifications as would be obvious to one skilled in the art are intended to be included within the scope of the following claims.

1. A control module for a provider edge node of a provider network, comprising:

a session table storing control information relating to peering sessions between the provider edge node and the provider network,

a circuit table storing, for each data tunnel, data tunnel identification data, encapsulation label data, outgoing

data packet interface identification data, CIR data, class data, setup priority data, and holding priority data;

a packet table storing, for each data packet flow, packet data interface identification data, data tunnel identification data, encapsulation label data, CIR data, class data, setup priority data, and holding priority data;

an ingress resource table storing, for each data tunnel, data tunnel identification data, physical bandwidth data and available bandwidth for each of a plurality of classes;

an egress resource table storing, for each egress data interface, egress data interface identification data, physical bandwidth data, available total bandwidth data, and available bandwidth for each of a plurality of classes;

admission control logic operatively connected to said session table, said circuit table, said packet table, said ingress resource table, and said egress resource table;

said admission control logic referring to said session table, said circuit table, said packet table, said ingress resource table, and said egress resource table to perform admission control on behalf of a new data flow requesting ingress to and egress from the provider network.

2. A method of establishing pseudo-wires between an initiating provider edge node and a terminating provider edge node of a provider network so as to permit admission control, comprising:

initiating a pseudo-wire request from the initiating provider edge node requesting a new data flow having a network resource requirement;

searching an ingress resource table for available network resources on one or more data tunnels to determine if there is a sufficient amount of available network resources on one or more data tunnels connecting the initiating and terminating provider edge nodes to satisfy the new data flow;

wherein upon the determination that there is a sufficient amount of available network resources to satisfy the new data flow, the method further comprises:

updating a packet table associated with the initiating provider edge node with the network resource requirement of the new data flow;

updating the ingress resource table with the network resources to be consumed by the new data flow; and

creating and sending a control message to the terminating provider edge node requesting a pseudo-wire to be set up between the initiating and terminating provider edge nodes to carry the new data flow.

3. The method according to claim 2, wherein at the terminating provider edge node the method further comprises:

receiving the control message requesting a pseudo-wire to be set up between the initiating and terminating provider edge nodes; and

determining if there are enough available network resources on the outgoing data interface to accommo-

date the new data flow based on the information contained in the control message and an egress resource table.

4. The method according to claim 3, wherein upon the determination that there is a sufficient amount of available network resources on the outgoing data interface to accommodate the new data flow, the method further comprises:

updating a circuit table associated with the terminating provider edge node with the network resource requirement data of the new data flow; and

updating an egress resource table with the network resources to be consumed by the new data flow.

5. The method according to claim 3, wherein upon the determination that there is not a sufficient amount of available network resources on the outgoing data interface to accommodate the new data flow, the method further comprises:

rejecting the pseudo-wire request.

6. The method according to claim 4,

wherein the network resource requirement of the new data flow includes CIR data and class data.

7. A method of establishing pseudo-wires between an initiating provider edge node and a terminating provider edge node of a provider network so as to permit admission control, comprising:

initiating a pseudo-wire request from the initiating provider edge node requesting a new data flow having network resource requirements;

determining if there is a sufficient amount of available network resources on one or more data tunnels connecting the initiating and terminating provider edge nodes to satisfy the new data flow; and

shuffling at least one existing pseudo-wire to accommodate the network resource requirements the new data flow if said determining step determines that there are insufficient available network resources.

8. The method according to claim 7, said shuffling further including exchanging control messages between the initiating and terminating provider edge nodes to adjust the shuffled pseudo-wires and ensure that the existing pseudo-wires are not disturbed due to said shuffling.

9. The method according to claim 7, further comprising:

preempting at least one existing pseudo-wire with the new data flow if said shuffling fails to provide enough network resources to accommodate the new data flow.

10. The method according to claim 9, said preempting further including exchanging control messages between the initiating and terminating provider edge nodes to adjust the at least one preempted pseudo-wire, ensure that the remaining pseudo-wires are not disturbed due to said preemption, and ensure that data flows associated with the preempted pseudo-wire can be properly terminated.

11. The method according to claim 9, further comprising:

re-determining, after said shuffling and said preempting, if there is a sufficient amount of available network resources on one or more data tunnels connecting the initiating and terminating provider edge nodes to satisfy the new data flow based on the network resource requirements of the new data flow and an ingress resource table.

12. The method according to claim 11, wherein upon the re-determination that there is not a sufficient amount of available network resources to satisfy the new data flow after said shuffling, the method further comprises:

rejecting the new data flow request.

13. The method according to claim 11, wherein upon the re-determination that there is a sufficient amount of available network resources to satisfy the new data flow after said shuffling, the method further comprises:

re-shuffling at least one existing pseudo-wire to accommodate the network resources of the new data flow.

14. The method according to claim 7,

said determining step searching an ingress resource table for available network resources on one or more data tunnels existing between the initiating and terminating provider edge nodes.

15. The method according to claim 14,

wherein upon the determination that there is a sufficient amount of available network resources to satisfy the new data flow, the method further comprises:

updating a packet table associated with the initiating provider edge node with the network resource requirements of the new data flow;

updating the ingress resource table with the network resources to be consumed by the new data flow; and

creating and sending a control message to the terminating provider edge node requesting a pseudo-wire to be set up within at least one data tunnel between the initiating and terminating provider edge nodes to carry the new data flow.

16. The method according to claim 15, wherein at the terminating provider edge node the method further comprises:

receiving the control message requesting a pseudo-wire to be set up between the initiating and terminating provider edge nodes; and

determining if there are enough available network resources on the outgoing data interface to accommodate the new data flow based on the information contained in the control message and an egress resource table.

17. The method according to claim 16, wherein if said determination step determines that there are not enough available network resources on the outgoing data interface to accommodate the new data flow, the method further comprises:

preempting bandwidth between the terminating provider edge node and a customer equipment node connected to the terminating provider edge node via the outgoing data interface and the associated pseudo-wire in order to accommodate the new data flow.

18. The method according to claim 17, further comprising:

re-determining, after said preempting, if there are enough available network resources on the outgoing data interface to accommodate the new data flow based on the information contained in the control message and an egress resource table.

19. The method according to claim 18, wherein upon the re-determination that there is not a sufficient amount of available network resources on the outgoing data interface to accommodate the new data flow, the method further comprises:

rejecting the pseudo-wire request.

20. The method according to claim 16, wherein upon the re-determination that there is not a sufficient amount of available network resources on the outgoing data interface to accommodate the new data flow, the method further comprises:

updating a circuit table associated with the terminating provider edge node with the network resource requirements of the new data flow; and

updating an egress resource table with the network resources to be consumed by the new data flow.

21. The method according to claim 20,

wherein the network resource requirements of the new data flow include CIR data, class data, setup priority data, and holding priority data.

\* \* \* \* \*

Network Working Group  
Request for Comments: 3386  
Category: Informational

W. Lai, Ed.  
AT&T  
D. McDysan, Ed.  
WorldCom  
November 2002

## Network Hierarchy and Multilayer Survivability

### Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

### Copyright Notice

Copyright (C) The Internet Society (2002). All Rights Reserved.

### Abstract

This document presents a proposal of the near-term and practical requirements for network survivability and hierarchy in current service provider environments.

### Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [2].



## Table of Contents

1. Introduction.....	2
2. Terminology and Concepts.....	5
2.1 Hierarchy.....	6
2.1.1 Vertical Hierarchy.....	5
2.1.2 Horizontal Hierarchy.....	6
2.2 Survivability Terminology.....	6
2.2.1 Survivability.....	7
2.2.2 Generic Operations.....	7
2.2.3 Survivability Techniques.....	8
2.2.4 Survivability Performance.....	9
2.3 Survivability Mechanisms: Comparison.....	10
3. Survivability.....	11
3.1 Scope.....	11
3.2 Required initial set of survivability mechanisms.....	12
3.2.1 1:1 Path Protection with Pre-Established Capacity.....	12
3.2.2 1:1 Path Protection with Pre-Planned Capacity.....	13
3.2.3 Local Restoration.....	13
3.2.4 Path Restoration.....	14
3.3 Applications Supported.....	14
3.4 Timing Bounds for Survivability Mechanisms.....	15
3.5 Coordination Among Layers.....	16
3.6 Evolution Toward IP Over Optical.....	17
4. Hierarchy Requirements.....	17
4.1 Historical Context.....	17
4.2 Applications for Horizontal Hierarchy.....	18
4.3 Horizontal Hierarchy Requirements.....	19
5. Survivability and Hierarchy.....	19
6. Security Considerations.....	20
7. References.....	21
8. Acknowledgments.....	22
9. Contributing Authors.....	22
Appendix A: Questions used to help develop requirements.....	23
Editors' Addresses.....	26
Full Copyright Statement.....	27

## 1. Introduction

This document is the result of the Network Hierarchy and Survivability Techniques Design Team established within the Traffic Engineering Working Group. This team collected and documented current and near term requirements for survivability and hierarchy in service provider environments. For clarity, an expanded set of definitions is included. The team determined that there appears to be a need to define a small set of interoperable survivability approaches in packet and non-packet networks. Suggested approaches include path-based as well as one that repairs connections in

proximity to the network fault. They operate primarily at a single network layer. For hierarchy, there did not appear to be a driving near-term need for work on "vertical hierarchy," defined as communication between network layers such as Time Division Multiplexed (TDM)/optical and Multi-Protocol Label Switching (MPLS). In particular, instead of direct exchange of signaling and routing between vertical layers, some looser form of coordination and communication, such as the specification of hold-off timers, is a nearer term need. For "horizontal hierarchy" in data networks, there are several pressing needs. The requirement is to be able to set up many Label Switched Paths (LSPs) in a service provider network with hierarchical Interior Gateway Protocol (IGP). This is necessary to support layer 2 and layer 3 Virtual Private Network (VPN) services that require edge-to-edge signaling across a core network.

This document presents a proposal of the near-term and practical requirements for network survivability and hierarchy in current service provider environments. With feedback from the working group solicited, the objective is to help focus the work that is being addressed in the TEWG (Traffic Engineering Working Group), CCAMP (Common Control and Measurement Plane Working Group), and other working groups. A main goal of this work is to provide some expedience for required functionality in multi-vendor service provider networks. The initial focus is primarily on intra-domain operations. However, to maintain consistency in the provision of end-to-end service in a multi-provider environment, rules governing the operations of survivability mechanisms at domain boundaries must also be specified. While such issues are raised and discussed, where appropriate, they will not be treated in depth in the initial release of this document.

The document first develops a set of definitions to be used later in this document and potentially in other documents as well. It then addresses the requirements and issues associated with service restoration, hierarchy, and finally a short discussion of survivability in hierarchical context.

Here is a summary of the findings:

#### A. Survivability Requirements

- o need to define a small set of interoperable survivability approaches in packet and non-packet networks
- o suggested survivability mechanisms include
  - 1:1 path protection with pre-established backup capacity (non-shared)
  - 1:1 path protection with pre-planned backup capacity (shared)

- local restoration with repairs in proximity to the network fault
- path restoration through source-based rerouting
- o timing bounds for service restoration to support voice call cutoff (140 msec to 2 sec), protocol timer requirements in premium data services, and mission critical applications
- o use of restoration priority for service differentiation

## B. Hierarchy Requirements

### B.1. Horizontally Oriented Hierarchy (Intra-Domain)

- o ability to set up many LSPs in a service provider network with hierarchical IGP, for the support of layer 2 and layer 3 VPN services
- o requirements for multi-area traffic engineering need to be developed to provide guidance for any necessary protocol extensions

### B.2. Vertically Oriented Hierarchy

The following functionality for survivability is common on most routing equipment today.

- o near-term need is some loose form of coordination and communication based on the use of nested hold-off timers, instead of direct exchange of signaling and routing between vertical layers
- o means for an upper layer to immediately begin recovery actions in the event that a lower layer is not configured to perform recovery

## C. Survivability Requirements in Horizontal Hierarchy

- o protection of end-to-end connection is based on a concatenated set of connections, each protected within their area
- o mechanisms for connection routing may include (1) a network element that participates on both sides of a boundary (e.g., OSPF ABR) - note that this is a common point of failure; (2) a route server
- o need for inter-area signaling of survivability information (1) to enable a "least common denominator" survivability mechanism at the boundary; (2) to convey the success or failure of the service restoration action; e.g., if a part of a "connection" is down on one side of a boundary, there is no need for the other side to recover from failures

## 2. Terminology and Concepts

### 2.1 Hierarchy

Hierarchy is a technique used to build scalable complex systems. It is based on an abstraction, at each level, of what is most significant from the details and internal structures of the levels further away. This approach makes use of a general property of all hierarchical systems composed of related subsystems that interactions between subsystems decrease as the level of communication between subsystems decreases.

Network hierarchy is an abstraction of part of a network's topology, routing and signaling mechanisms. Abstraction may be used as a mechanism to build large networks or as a technique for enforcing administrative, topological, or geographic boundaries. For example, network hierarchy might be used to separate the metropolitan and long-haul regions of a network, or to separate the regional and backbone sections of a network, or to interconnect service provider networks (with BGP which reduces a network to an Autonomous System).

In this document, network hierarchy is considered from two perspectives:

- (1) Vertically oriented: between two network technology layers.
- (2) Horizontally oriented: between two areas or administrative subdivisions within the same network technology layer.

#### 2.1.1 Vertical Hierarchy

Vertical hierarchy is the abstraction, or reduction in information, which would be of benefit when communicating information across network technology layers, as in propagating information between optical and router networks.

In the vertical hierarchy, the total network functions are partitioned into a series of functional or technological layers with clear logical, and maybe even physical, separation between adjacent layers. Survivability mechanisms either currently exist or are being developed at multiple layers in networks [3]. The optical layer is now becoming capable of providing dynamic ring and mesh restoration functionality, in addition to traditional 1+1 or 1:1 protection. The Synchronous Digital Hierarchy (SDH)/Synchronous Optical NETWORK (SONET) layer provides survivability capability with automatic protection switching (APS), as well as self-healing ring and mesh restoration architectures. Similar functionality has been defined in the Asynchronous Transfer Mode (ATM) Layer, with work ongoing to also provide such functionality using MPLS [4]. At the IP layer,

rerouting is used to restore service continuity following link and node outages. Rerouting at the IP layer, however, occurs after a period of routing convergence, which may require a few seconds to several minutes to complete [5].

### 2.1.2 Horizontal Hierarchy

Horizontal hierarchy is the abstraction that allows a network at one technology layer, for instance a packet network, to scale. Examples of horizontal hierarchy include BGP confederations, separate Autonomous Systems, and multi-area OSPF.

In the horizontal hierarchy, a large network is partitioned into multiple smaller, non-overlapping sub-networks. The partitioning criteria can be based on topology, network function, administrative policy, or service domain demarcation. Two networks at the \*same\* hierarchical level, e.g., two Autonomous Systems in BGP, may share a peer relation with each other through some loose form of coupling. On the other hand, for routing in large networks using multi-area OSPF, abstraction through the aggregation of routing information is achieved through a hierarchical partitioning of the network.

## 2.2 Survivability Terminology

In alphabetical order, the following terms are defined in this section:

- backup entity, same as protection entity (section 2.2.2)
- extra traffic (section 2.2.2)
- non-revertive mode (section 2.2.2)
- normalization (section 2.2.2)
- preemptable traffic, same as extra traffic (section 2.2.2)
- preemption priority (section 2.2.4)
- protection (section 2.2.3)
- protection entity (section 2.2.2)
- protection switching (section 2.2.3)
- protection switch time (section 2.2.4)
- recovery (section 2.2.2)
- recovery by rerouting, same as restoration (section 2.2.3)
- recovery entity, same as protection entity (section 2.2.2)
- restoration (section 2.2.3)
- restoration priority (section 2.2.4)
- restoration time (section 2.2.4)
- revertive mode (section 2.2.2)
- shared risk group (SRG) (section 2.2.2)
- survivability (section 2.2.1)
- working entity (section 2.2.2)

### 2.2.1 Survivability

Survivability is the capability of a network to maintain service continuity in the presence of faults within the network [6]. Survivability mechanisms such as protection and restoration are implemented either on a per-link basis, on a per-path basis, or throughout an entire network to alleviate service disruption at affordable costs. The degree of survivability is determined by the network's capability to survive single failures, multiple failures, and equipment failures.

### 2.2.2 Generic Operations

This document does not discuss the sequence of events of how network failures are monitored, detected, and mitigated. For more detail of this aspect, see [4]. Also, the repair process following a failure is out of the scope here.

A working entity is the entity that is used to carry traffic in normal operation mode. Depending upon the context, an entity can be a channel or a transmission link in the physical layer, an Label Switched Path (LSP) in MPLS, or a logical bundle of one or more LSPs.

A protection entity, also called backup entity or recovery entity, is the entity that is used to carry protected traffic in recovery operation mode, i.e., when the working entity is in error or has failed.

Extra traffic, also referred to as preemptable traffic, is the traffic carried over the protection entity while the working entity is active. Extra traffic is not protected, i.e., when the protection entity is required to protect the traffic that is being carried over the working entity, the extra traffic is preempted.

A shared risk group (SRG) is a set of network elements that are collectively impacted by a specific fault or fault type. For example, a shared risk link group (SRLG) is the union of all the links on those fibers that are routed in the same physical conduit in a fiber-span network. This concept includes, besides shared conduit, other types of compromise such as shared fiber cable, shared right of way, shared optical ring, shared office without power sharing, etc. The span of an SRG, such as the length of the sharing for compromised outside plant, needs to be considered on a per fault basis. The concept of SRG can be extended to represent a "risk domain" and its associated capabilities and summarization for traffic engineering purposes. See [7] for further discussion.

Normalization is the sequence of events and actions taken by a network that returns the network to the preferred state upon completing repair of a failure. This could include the switching or rerouting of affected traffic to the original repaired working entities or new routes. Revertive mode refers to the case where traffic is automatically returned to a repaired working entity (also called switch back).

Recovery is the sequence of events and actions taken by a network after the detection of a failure to maintain the required performance level for existing services (e.g., according to service level agreements) and to allow normalization of the network. The actions include notification of the failure followed by two parallel processes: (1) a repair process with fault isolation and repair of the failed components, and (2) a reconfiguration process using survivability mechanisms to maintain service continuity. In protection, reconfiguration involves switching the affected traffic from a working entity to a protection entity. In restoration, reconfiguration involves path selection and rerouting for the affected traffic.

Revertive mode is a procedure in which revertive action, i.e., switch back from the protection entity to the working entity, is taken once the failed working entity has been repaired. In non-revertive mode, such action is not taken. To minimize service interruption, switch-back in revertive mode should be performed at a time when there is the least impact on the traffic concerned, or by using the make-before-break concept.

Non-revertive mode is the case where there is no preferred path or it may be desirable to minimize further disruption of the service brought on by a revertive switching operation. A switch-back to the original working path is not desired or not possible since the original path may no longer exist after the occurrence of a fault on that path.

### 2.2.3 Survivability Techniques

Protection, also called protection switching, is a survivability technique based on predetermined failure recovery: as the working entity is established, a protection entity is also established. Protection techniques can be implemented by several architectures: 1+1, 1:1, 1:n, and m:n. In the context of SDH/SONET, they are referred to as Automatic Protection Switching (APS).

In the 1+1 protection architecture, a protection entity is dedicated to each working entity. The dual-feed mechanism is used whereby the working entity is permanently bridged onto the protection entity at

the source of the protected domain. In normal operation mode, identical traffic is transmitted simultaneously on both the working and protection entities. At the other end (sink) of the protected domain, both feeds are monitored for alarms and maintenance signals. A selection between the working and protection entity is made based on some predetermined criteria, such as the transmission performance requirements or defect indication.

In the 1:1 protection architecture, a protection entity is also dedicated to each working entity. The protected traffic is normally transmitted by the working entity. When the working entity fails, the protected traffic is switched to the protection entity. The two ends of the protected domain must signal detection of the fault and initiate the switchover.

In the 1:n protection architecture, a dedicated protection entity is shared by n working entities. In this case, not all of the affected traffic may be protected.

The m:n architecture is a generalization of the 1:n architecture. Typically  $m \leq n$ , where m dedicated protection entities are shared by n working entities.

Restoration, also referred to as recovery by rerouting [4], is a survivability technique that establishes new paths or path segments on demand, for restoring affected traffic after the occurrence of a fault. The resources in these alternate paths are the currently unassigned (unreserved) resources in the same layer. Preemption of extra traffic may also be used if spare resources are not available to carry the higher-priority protected traffic. As initiated by detection of a fault on the working path, the selection of a recovery path may be based on preplanned configurations, network routing policies, or current network status such as network topology and fault information. Signaling is used for establishing the new paths to bypass the fault. Thus, restoration involves a path selection process followed by rerouting of the affected traffic from the working entity to the recovery entity.

#### 2.2.4 Survivability Performance

Protection switch time is the time interval from the occurrence of a network fault until the completion of the protection-switching operations. It includes the detection time necessary to initiate the protection switch, any hold-off time to allow for the interworking of protection schemes, and the switch completion time.



Restoration time is the time interval from the occurrence of a network fault to the instant when the affected traffic is either completely restored, or until spare resources are exhausted, and/or no more extra traffic exists that can be preempted to make room.

Restoration priority is a method of giving preference to protect higher-priority traffic ahead of lower-priority traffic. Its use is to help determine the order of restoring traffic after a failure has occurred. The purpose is to differentiate service restoration time as well as to control access to available spare capacity for different classes of traffic.

Preemption priority is a method of determining which traffic can be disconnected in the event that not all traffic with a higher restoration priority is restored after the occurrence of a failure.

### 2.3 Survivability Mechanisms: Comparison

In a survivable network design, spare capacity and diversity must be built into the network from the beginning to support some degree of self-healing whenever failures occur. A common strategy is to associate each working entity with a protection entity having either dedicated resources or shared resources that are pre-reserved or reserved-on-demand. According to the methods of setting up a protection entity, different approaches to providing survivability can be classified. Generally, protection techniques are based on having a dedicated protection entity set up prior to failure. Such is not the case in restoration techniques, which mainly rely on the use of spare capacity in the network. Hence, in terms of trade-offs, protection techniques usually offer fast recovery from failure with enhanced availability, while restoration techniques usually achieve better resource utilization.

A 1+1 protection architecture is rather expensive since resource duplication is required for the working and protection entities. It is generally used for specific services that need a very high availability.

A 1:1 architecture is inherently slower in recovering from failure than a 1+1 architecture since communication between both ends of the protection domain is required to perform the switch-over operation. An advantage is that the protection entity can optionally be used to carry low-priority extra traffic in normal operation, if traffic preemption is allowed. Packet networks can pre-establish a protection path for later use with pre-planned but not pre-reserved capacity. That is, if no packets are sent onto a protection path,

then no bandwidth is consumed. This is not the case in transmission networks like optical or TDM where path establishment and resource reservation cannot be decoupled.

In the 1:n protection architecture, traffic is normally sent on the working entities. When multiple working entities have failed simultaneously, only one of them can be restored by the common protection entity. This contention could be resolved by assigning a different preemptive priority to each working entity. As in the 1:1 case, the protection entity can optionally be used to carry preemptable traffic in normal operation.

While the m:n architecture can improve system availability with small cost increases, it has rarely been implemented or standardized.

When compared with protection mechanisms, restoration mechanisms are generally more frugal as no resources are committed until after the fault occurs and the location of the fault is known. However, restoration mechanisms are inherently slower, since more must be done following the detection of a fault. Also, the time it takes for the dynamic selection and establishment of alternate paths may vary, depending on the amount of traffic and connections to be restored, and is influenced by the network topology, technology employed, and the type and severity of the fault. As a result, restoration time tends to be more variable than the protection switch time needed with pre-selected protection entities. Hence, in using restoration mechanisms, it is essential to use restoration priority to ensure that service objectives are met cost-effectively.

Once the network routing algorithms have converged after a fault, it may be preferable in some cases, to reoptimize the network by performing a reroute based on the current state of the network and network policies.

### 3. Survivability

#### 3.1 Scope

Interoperable approaches to network survivability were determined to be an immediate requirement in packet networks as well as in SDH/SONET framed TDM networks. Not as pressing at this time were techniques that would cover all-optical networks (e.g., where framing is unknown), as the control of these networks in a multi-vendor environment appeared to have some other hurdles to first deal with. Also, not of immediate interest were approaches to coordinate or explicitly communicate survivability mechanisms across network layers (such as from a TDM or optical network to/from an IP network). However, a capability should be provided for a network operator to

perform fault notification and to control the operation of survivability mechanisms among different layers. This may require the development of corresponding OAM functionality. However, such issues and those related to OAM are currently outside the scope of this document. (For proposed MPLS OAM requirements, see [8, 9]).

The initial scope is to address only "backhoe failures" in the inter-office connections of a service provider network. A link connection in the router layer is typically comprised of multiple spans in the lower layers. Therefore, the types of network failures that cause a recovery to be performed include link/span failures. However, linecard and node failures may not need to be treated any differently than their respective link/span failures, as a router failure may be represented as a set of simultaneous link failures.

Depending on the actual network configuration, drop-side interface (e.g., between a customer and an access router, or between a router and an optical cross-connect) may be considered either inter-domain or inter-layer. Another inter-domain scenario is the use of intra-office links for interconnecting a metro network and a core network, with both networks being administered by the same service provider. Failures at such interfaces may be similarly protected by the mechanisms of this section.

Other more complex failure mechanisms such as systematic control-plane failure, configuration error, or breach of security are not within the scope of the survivability mechanisms discussed in this document. Network impairment such as congestion that results in lower throughput are also not covered.

### 3.2 Required initial set of survivability mechanisms

#### 3.2.1 1:1 Path Protection with Pre-Established Capacity

In this protection mode, the head end of a working connection establishes a protection connection to the destination. There should be the ability to maintain relative restoration priorities between working and protection connections, as well as between different classes of protection connections.

In normal operation, traffic is only sent on the working connection, though the ability to signal that traffic will be sent on both connections (1+1 Path for signaling purposes) would be valuable in non-packet networks. Some distinction between working and protection connections is likely, either through explicit objects, or preferably through implicit methods such as general classes or priorities. Head ends need the ability to create connections that are as failure disjoint as possible from each other. This requires SRG information

that can be generally assigned to either nodes or links and propagated through the control or management plane. In this mechanism, capacity in the protection connection is pre-established, however it should be capable of carrying preemptable extra traffic in non-packet networks. When protection capacity is called into service during recovery, there should be the ability to promote the protection connection to working status (for non-revertive mode operation) with some form of make-before-break capability.

### 3.2.2 1:1 Path Protection with Pre-Planned Capacity

Similar to the above 1:1 protection with pre-established capacity, the protection connection in this case is also pre-sigaled. The difference is in the way protection capacity is assigned. With pre-planned capacity, the mechanism supports the ability for the protection capacity to be shared, or "double-booked". Operators need the ability to provision different amounts of protection capacity according to expected failure modes and service level agreements. Thus, an operator may wish to provision sufficient restoration capacity to handle a single failure affecting all connections in an SRG, or may wish to provision less or more restoration capacity. Mechanisms should be provided to allow restoration capacity on each link to be shared by SRG-disjoint failures. In a sense, this is 1:1 from a path perspective; however, the protection capacity in the network (on a link by link basis) is shared in a 1:n fashion, e.g., see the proposals in [10, 11]. If capacity is planned but not allocated, some form of signaling could be required before traffic may be sent on protection connections, especially in TDM networks.

The use of this approach improves network resource utilization, but may require more careful planning. So, initial deployment might be based on 1:1 path protection with pre-established capacity and the local restoration mechanism to be described next.

### 3.2.3 Local Restoration

Due to the time impact of signal propagation, dynamic recovery of an entire path may not meet the service requirements of some networks. The solution to this is to restore connectivity of the link or span in immediate proximity to the fault, e.g., see the proposals in [12, 13]. At a minimum, this approach should be able to protect against connectivity-type SRGs, though protecting against node-based SRGs might be worthwhile. Also, this approach is applicable to support restoration on the inter-domain and inter-layer interconnection scenarios using intra-office links as described in the Scope Section.

Head end systems must have some control as to whether their connections are candidates for or excluded from local restoration. For example, best-effort and preemptable traffic may be excluded from local restoration; they only get restored if there is bandwidth available. This type of control may require the definition of an object in signaling.

Since local restoration may be suboptimal, a means for head end systems to later perform path-level re-grooming must be supported for this approach.

#### 3.2.4 Path Restoration

In this approach, connections that are impacted by a fault are rerouted by the originating network element upon notification of connection failure. Such a source-based approach is efficient for network resources, but typically takes longer to accomplish restoration. It does not involve any new mechanisms. It merely is a mention of another common approach to protecting against faults in a network.

#### 3.3 Applications Supported

With service continuity under failure as a goal, a network is "survivable" if, in the face of a network failure, connectivity is interrupted for a "brief" period and then recovered before the network failure ends. The length of this interrupted period is dependent upon the application supported. Here are some typical applications and considerations that drive the requirements for an acceptable protection switch time or restoration time:

- Best-effort data: recovery of network connectivity by rerouting at the IP layer would be sufficient
- Premium data service: need to meet TCP timeout or application protocol timer requirements
- Voice: call cutoff is in the range of 140 msec to 2 sec (the time that a person waits after interruption of the speech path before hanging up or the time that a telephone switch will disconnect a call)
- Other real-time service (e.g., streaming, fax) where an interruption would cause the session to terminate
- Mission-critical applications that cannot tolerate even brief interruptions, for example, real-time financial transactions

### 3.4 Timing Bounds for Survivability Mechanisms

The approach to picking the types of survivability mechanisms recommended was to consider a spectrum of mechanisms that can be used to protect traffic with varying characteristics of survivability and speed of protection/restoration, and then attempt to select a few general points that provide some coverage across that spectrum. The focus of this work is to provide requirements to which a small set of detailed proposals may be developed, allowing the operator some (limited) flexibility in approaches to meeting their design goals in engineering multi-vendor networks. Requirements of different applications as listed in the previous sub-section were discussed generally, however none on the team would likely attest to the scientific merit of the ability of the timing bounds below to meet any specific application's needs. A few assumptions include:

1. Approaches in which protection switch without propagation of information are likely to be faster than those that do require some form of fault notification to some or all elements in a network.
2. Approaches that require some form of signaling after a fault will also likely suffer some timing impact.

Proposed timing bounds for different survivability mechanisms are as follows (all bounds are exclusive of signal propagation):

1:1 path protection with pre-established capacity:	100-500 ms
1:1 path protection with pre-planned capacity:	100-750 ms
Local restoration:	50 ms
Path restoration:	1-5 seconds

To ensure that the service requirements for different applications can be met within the above timing bounds, restoration priority must be implemented to determine the order in which connections are restored (to minimize service restoration time as well as to gain access to available spare capacity on the best paths). For example, mission critical applications may require high restoration priority. At the fiber layer, instead of specific applications, it may be possible that priority be given to certain classifications of customers with their traffic types enclosed within the customer aggregate. Preemption priority should only be used in the event that not all connections can be restored, in which case connections with lower preemption priority should be released. Depending on a service provider's strategy in provisioning network resources for backup, preemption may or may not be needed in the network.

### 3.5 Coordination Among Layers

A common design goal for networks with multiple technological layers is to provide the desired level of service in the most cost-effective manner. Multilayer survivability may allow the optimization of spare resources through the improvement of resource utilization by sharing spare capacity across different layers, though further investigations are needed. Coordination during recovery among different network layers (e.g., IP, SDH/SONET, optical layer) might necessitate development of vertical hierarchy. The benefits of providing survivability mechanisms at multiple layers, and the optimization of the overall approach, must be weighed with the associated cost and service impacts.

A default coordination mechanism for inter-layer interaction could be the use of nested timers and current SDH/SONET fault monitoring, as has been done traditionally for backward compatibility. Thus, when lower-layer recovery happens in a longer time period than higher-layer recovery, a hold-off timer is utilized to avoid contention between the different single-layer survivability schemes. In other words, multilayer interaction is addressed by having successively higher multiplexing levels operate at a protection/restoration time scale greater than the next lowest layer. This can impact the overall time to recover service. For example, if SDH/SONET protection switching is used, MPLS recovery timers must wait until SDH/SONET has had time to switch. Setting such timers involves a tradeoff between rapid recovery and creation of a race condition where multiple layers are responding to the same fault, potentially allocating resources in an inefficient manner.

In other configurations where the lower layer does not have a restoration capability or is not expected to protect, say an unprotected SDH/SONET linear circuit, then there must be a mechanism for the lower layer to trigger the higher layer to take recovery actions immediately. This difference in network configuration means that implementations must allow for adjustment of hold-off timer values and/or a means for a lower layer to immediately indicate to a higher layer that a fault has occurred so that the higher layer can take restoration or protection actions.

Furthermore, faults at higher layers should not trigger restoration or protection actions at lower layers [3, 4].

It was felt that the current approach to coordination of survivability approaches currently did not have significant operational shortfalls. These approaches include protecting traffic solely at one layer (e.g., at the IP layer over linear WDM, or at the SDH/SONET layer). Where survivability mechanisms might be deployed

at several layers, such as when a routed network rides a SDH/SONET protected network, it was felt that current coordination approaches were sufficient in many cases. One exception is the hold-off of MPLS recovery until the completion of SDH/SONET protection switching as described above. This limits the recovery time of fast MPLS restoration. Also, by design, the operations and mechanisms within a given layer tend to be invisible to other layers.

### 3.6 Evolution Toward IP Over Optical

As more pressing requirements for survivability and horizontal hierarchy for edge-to-edge signaling are met with technical proposals, it is believed that the benefits of merging (in some manner) the control planes of multiple layers will be outlined. When these benefits are self-evident, it would then seem to be the right time to review whether vertical hierarchy mechanisms are needed, and what the requirements might be. For example, a future requirement might be to provide a better match between the recovery requirements of IP networks with the recovery capability of optical transport. One such proposal is described in [14].

## 4. Hierarchy Requirements

Efforts in the area of network hierarchy should focus on mechanisms that would allow more scalable edge-to-edge signaling, or signaling across networks with existing network hierarchy (such as multi-area OSPF). This appears to be a more urgent need than mechanisms that might be needed to interconnect networks at different layers.

### 4.1 Historical Context

One reason for horizontal hierarchy is functionality (e.g., metro versus backbone). Geographic "islands" or partitions reduce the need for interoperability and make administration and operations less complex. Using a simpler, more interoperable, survivability scheme at metro/backbone boundaries is natural for many provider network architectures. In transmission networks, creating geographic islands of different vendor equipment has been done for a long time because multi-vendor interoperability has been difficult to achieve. Traditionally, providers have to coordinate the equipment on either end of a "connection," and making this interoperable reduces complexity. A provider should be able to concatenate survivability mechanisms in order to provide a "protected link" to the next higher level. Think of SDH/SONET rings connecting to TDM DXCs with 1+1 line-layer protection between the ADM and the DXC port. The TDM connection, e.g., a DS3, is protected but usually all equipment on each SDH/SONET ring is from a single vendor. The DXC cross connections are controlled by the provider and the ports are



physically protected resulting in a highly available design. Thus, concatenation of survivability approaches can be used to cascade across a horizontal hierarchy. While not perfect, it is workable in the near to mid-term until multi-vendor interoperability is achieved.

While the problems associated with multi-vendor interoperability may necessitate horizontal hierarchy as a practical matter in the near to mid-term (at least this has been the case in TDM networks), there should not be a technical reason for it in the standards developed by the IETF for core networks, or even most access networks. Establishing interoperability of survivability mechanisms between multi-vendor equipment in core IP networks is urgently required to enable adoption of IP as a viable core transport technology and to facilitate the traffic engineering of future multi-service IP networks [3].

Some of the largest service provider networks currently run a single area/level IGP. Some service providers, as well as many large enterprise networks, run multi-area Open Shortest Path First (OSPF) to gain increases in scalability. Often, this was from an original design, so it is difficult to say if the network truly required the hierarchy to reach its current size.

Some proposals on improved mechanisms to address network hierarchy have been suggested [15, 16, 17, 18, 19]. This document aims to provide the concrete requirements so that these and other proposals can first aim to meet some limited objectives.

#### 4.2 Applications for Horizontal Hierarchy

A primary driver for intra-domain horizontal hierarchy is signaling capabilities in the context of edge-to-edge VPNs, potentially across traffic-engineered data networks. There are a number of different approaches to layer 2 and layer 3 VPNs and they are currently being addressed by different emerging protocols in the provider-provisioned VPNs (e.g., virtual routers) and Pseudo Wire Edge-to-Edge Emulation (PWE3) efforts based on either MPLS and/or IP tunnels. These may or may not need explicit signaling from edge to edge, but it is a common perception that in order to meet SLAs, some form of edge-to-edge signaling may be required.

With a large number of edges ( $N$ ), scalability is concerned with avoiding the  $O(N^2)$  properties of edge-to-edge signaling. However, the main issue here is not with the scalability of large amounts of signaling, such as in  $O(N^2)$  meshes with a "connection" between every edge-pair. This is because, even if establishing and maintaining connections is feasible in a large network, there might be an impact on core survivability mechanisms which would cause

protection/restoration times to grow with  $N^2$ , which would be undesirable. While some value of  $N$  may be inevitable, approaches to reduce  $N$  (e.g. to pull in from the edge to aggregation points) might be of value.

Thus, most service providers feel that  $O(N^2)$  meshes are not necessary for VPNs, and that the number of tunnels to support VPNs would be within the scalability bounds of current protocols and implementations. That may be the case, as there is currently a lack of ability to signal MPLS tunnels from edge to edge across IGP hierarchy, such as OSPF areas. This may require the development of signaling standards that support dynamic establishment and potentially the restoration of LSPs across a 2-level IGP hierarchy.

For routing scalability, especially in data applications, a major concern is the amount of processing/state that is required in the variety of network elements. If some nodes might not be able to communicate and process the state of every other node, it might be preferable to limit the information. There is one school of thought that says that the amount of information contained by a horizontal barrier should be significant, and that impacts this might have on optimality in route selection and ability to provide global survivability are accepted tradeoffs.

#### 4.3 Horizontal Hierarchy Requirements

Mechanisms are required to allow for edge-to-edge signaling of connections through a network. One network scenario includes medium to large networks that currently have hierarchical interior routing such as multi-area OSPF or multi-level Intermediate System to Intermediate System (IS-IS). The primary context of this is edge-to-edge signaling, which is thought to be required to assure the SLAs for the layer 2 and layer 3 VPNs that are being carried across the network. Another possible context would be edge-to-edge signaling in TDM SDH/SONET networks with IP control, where metro and core networks again might be in a hierarchical interior routing domain.

To support edge-to-edge signaling in the above network scenarios within the framework of existing horizontal hierarchies, current traffic engineering (TE) methods [20, 6] may need to be extended. Requirements for multi-area TE need to be developed to provide guidance for any necessary protocol extensions.

#### 5. Survivability and Hierarchy

When horizontal hierarchy exists in a network technology layer, a question arises as to how survivability can be provided along a connection that crosses hierarchical boundaries.

In designing protocols to meet the requirements of hierarchy, an approach to consider is that boundaries are either clean, or are of minimal value. However, the concept of network elements that participate on both sides of a boundary might be a consideration (e.g., OSPF ABRs). That would allow for devices on either side to take an intra-area approach within their region of knowledge, and for the ABR to do this in both areas, and splice the two protected connections together at a common point (granted it is a common point of failure now). If the limitations of this approach start to appear in operational settings, then perhaps it would be time to start thinking about route-servers and signaling propagated directives. However, one initial approach might be to signal through a common border router, and to consider the service as protected as it consists of a concatenated set of connections which are each protected within their area. Another approach might be to have a least common denominator mechanism at the boundary, e.g., 1+1 port protection. There should also be some standardized means for a survivability scheme on one side of such a boundary to communicate with the scheme on the other side regarding the success or failure of the recovery action. For example, if a part of a "connection" is down on one side of such a boundary, there is no need for the other side to recover from failures.

In summary, at this time, approaches as described above that allow concatenation of survivability schemes across hierarchical boundaries seem sufficient.

## 6. Security Considerations

The set of SRGs that are defined for a network under a common administrative control and the corresponding assignment of these SRGs to nodes and links within the administrative control is sensitive information and needs to be protected. An SRG is an acknowledgement that nodes and links that belong to an SRG are susceptible to a common threat. An adversary with access to information contained in an SRG could use that information to design an attack, determine the scope of damage caused by the attack and, therefore, be used to maximize the effect of an attack.

The label used to refer to a particular SRG must allow for an encoding such that sensitive information such as physical location, function, purpose, customer, fault type, etc. is not readily discernable by unauthorized users.

SRG information that is propagated through the control and management plane should allow for an encryption mechanism. An example of an approach would be to use IPSEC [21] on all packets carrying SRG information.

## 7. References

- [1] Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.
- [2] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [3] K. Owens, V. Sharma, and M. Oommen, "Network Survivability Considerations for Traffic Engineered IP Networks", Work in Progress.
- [4] V. Sharma, B. Crane, S. Makam, K. Owens, C. Huang, F. Hellstrand, J. Weil, L. Andersson, B. Jamoussi, B. Cain, S. Civanlar, and A. Chiu, "Framework for MPLS-based Recovery", Work in Progress.
- [5] M. Thorup, "Fortifying OSPF/ISIS Against Link Failure", [http://www.research.att.com/~mthorup/PAPERS/lf\\_ospf.ps](http://www.research.att.com/~mthorup/PAPERS/lf_ospf.ps)
- [6] Awduche, D., Chiu, A., Elwalid, A., Widjaja, I. and X. Xiao, "Overview and Principles of Internet Traffic Engineering", RFC 3272, May 2002.
- [7] S. Dharanikota, R. Jain, D. Papadimitriou, R. Hartani, G. Bernstein, V. Sharma, C. Brownmiller, Y. Xue, and J. Strand, "Inter-domain routing with Shared Risk Groups", Work in Progress.
- [8] N. Harrison, P. Willis, S. Davari, E. Cuevas, B. Mack-Crane, E. Franze, H. Ohta, T. So, S. Goldfless, and F. Chen, "Requirements for OAM in MPLS Networks," Work in Progress.
- [9] D. Allan and M. Azad, "A Framework for MPLS User Plane OAM," Work in Progress.
- [10] S. Kini, M. Kodialam, T.V. Lakshman, S. Sengupta, and C. Villamizar, "Shared Backup Label Switched Path Restoration," Work in Progress.
- [11] G. Li, C. Kalmanek, J. Yates, G. Bernstein, F. Liaw, and V. Sharma, "RSVP-TE Extensions For Shared-Mesh Restoration in Transport Networks", Work in Progress.
- [12] P. Pan (Editor), D.H. Gan, G. Swallow, J. Vasseur, D. Cooper, A. Atlas, and M. Jork, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", Work in Progress.

- [13] A. Atlas, C. Villamizar, and C. Litvanyi, "MPLS RSVP-TE Interoperability for Local Protection/Fast Reroute", Work in Progress.
- [14] A. Chiu and J. Strand, "Joint IP/Optical Layer Restoration after a Router Failure", Proc. OFC'2001, Anaheim, CA, March 2001.
- [15] K. Kompella and Y. Rekhter, "Multi-area MPLS Traffic Engineering", Work in Progress.
- [16] G. Ash, et. al., "Requirements for Multi-Area TE", Work in Progress.
- [17] A. Iwata, N. Fujita, G.R. Ash, and A. Farrel, "Crankback Routing Extensions for MPLS Signaling", Work in Progress.
- [18] C-Y Lee, A. Celer, N. Gammage, S. Ghanti, G. Ash, "Distributed Route Exchangers", Work in Progress.
- [19] C-Y Lee and S. Ghanti, "Path Request and Path Reply Message", Work in Progress.
- [20] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M. and J. McManus, "Requirements for Traffic Engineering Over MPLS", RFC 2702, September 1999.
- [21] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.

## 8. Acknowledgments

A lot of the direction taken in this document, and by the team in its initial effort was steered by the insightful questions provided by Bala Rajagoplan, Greg Bernstein, Yangguang Xu, and Avri Doria. The set of questions is attached as Appendix A in this document.

After the release of the first draft, a number of comments were received. Thanks to the inputs from Jerry Ash, Sudheer Dharanikota, Chuck Kalmanek, Dan Koller, Lyndon Ong, Steve Plote, and Yong Xue.

## 9. Contributing Authors

Jim Boyle (PDNets), Rob Coltun (Movaz), Tim Griffin (AT&T), Ed Kern, Tom Reddington (Lucent) and Malin Carlzon.

## Appendix A: Questions used to help develop requirements

## A. Definitions

1. In determining the specific requirements, the design team should precisely define the concepts "survivability", "restoration", "protection", "protection switching", "recovery", "re-routing" etc. and their relations. This would enable the requirements doc to describe precisely which of these will be addressed. In the following, the term "restoration" is used to indicate the broad set of policies and mechanisms used to ensure survivability.

## B. Network types and protection modes

1. What is the scope of the requirements with regard to the types of networks covered? Specifically, are the following in scope:

Restoration of connections in mesh optical networks (opaque or transparent)

Restoration of connections in hybrid mesh-ring networks

Restoration of LSPs in MPLS networks (composed of LSRs overlaid on a transport network, e.g., optical)

Any other types of networks?

Is commonality of approach, or optimization of approach more important?

2. What are the requirements with regard to the protection modes to be supported in each network type covered? (Examples of protection modes include 1+1, M:N, shared mesh, UPSR, BLSR, newly defined modes such as P-cycles, etc.)
3. What are the requirements on local span (i.e., link by link) protection and end-to-end protection, and the interaction between them? E.g.: what should be the granularity of connections for each type (single connection, bundle of connections, etc).

## C. Hierarchy

1. Vertical (between two network layers):  
What are the requirements for the interaction between restoration procedures across two network layers, when these features are offered in both layers? (Example, MPLS network realized over pt-to-pt optical connections.) Under such a case,
  - (a) Are there any criteria to choose which layer should provide protection?

- (b) If both layers provide survivability features, what are the requirements to coordinate these mechanisms?
  - (c) How is lack of current functionality of cross-layer coordination currently hampering operations?
  - (d) Would the benefits be worth additional complexity associated with routing isolation (e.g. VPN, areas), security, address isolation and policy / authentication processes?
2. Horizontal (between two areas or administrative subdivisions within the same network layer):
- (a) What are the criteria that trigger the creation of protocol or administrative boundaries pertaining to restoration? (e.g., scalability? multi-vendor interoperability? what are the practical issues?) multi-provider? Should multi-vendor necessitate hierarchical separation?

When such boundaries are defined:

- (b) What are the requirements on how protection/restoration is performed end-to-end across such boundaries?
- (c) If different restoration mechanisms are implemented on two sides of a boundary, what are the requirements on their interaction?

What is the primary driver of horizontal hierarchy? (select one)

- functionality (e.g. metro -v- backbone)
- routing scalability
- signaling scalability
- current network architecture, trying to layer on TE on top of an already hierarchical network architecture
- routing and signalling

For signalling scalability, is it

- manageability
- processing/state of network
- edge-to-edge N<sup>2</sup> type issue

For routing scalability, is it

- processing/state of network
- are you flat and want to go hierarchical
- or already hierarchical?
- data or TDM application?

## D. Policy

1. What are the requirements for policy support during protection/restoration, e.g., restoration priority, preemption, etc.

## E. Signaling Mechanisms

1. What are the requirements on the signaling transport mechanism (e.g., in-band over SDH/SONET overhead bytes, out-of-band over an IP network, etc.) used to communicate restoration protocol messages between network elements? What are the bandwidth and other requirements on the signaling channels?
2. What are the requirements on fault detection/localization mechanisms (which is the prelude to performing restoration procedures) in the case of opaque and transparent optical networks? What are the requirements in the case of MPLS restoration?
3. What are the requirements on signaling protocols to be used in restoration procedures (e.g., high priority processing, security, etc)?
4. Are there any requirements on the operation of restoration protocols?

## F. Quantitative

1. What are the quantitative requirements (e.g., latency) for completing restoration under different protection modes (for both local and end-to-end protection)?

## G. Management

1. What information should be measured/maintained by the control plane at each network element pertaining to restoration events?
2. What are the requirements for the correlation between control plane and data plane failures from the restoration point of view?



Editors' Addresses

Wai Sum Lai  
AT&T  
200 Laurel Avenue  
Middletown, NJ 07748, USA

Phone: +1 732-420-3712  
EMail: wlai@att.com

Dave McDysan  
WorldCom  
22001 Loudoun County Pkwy  
Ashburn, VA 20147, USA

EMail: dave.mcdysan@wcom.com

## Full Copyright Statement

Copyright (C) The Internet Society (2002). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

## Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.





## Metro Ethernet

The definitive guide to enterprise and carrier  
metro Ethernet applications

JUNIPER Exhibit 1003  
App. 6, pg. 1



# Metro Ethernet

**Sam Halabi**

**Cisco Press**

800 East 96th Street, 3rd Floor  
Indianapolis, IN 46240 USA

JUNIPER Exhibit 1003  
App. 6, pg. 2

From the Library of Tal Lavian

## **Metro Ethernet**

Sam Halabi

Copyright© 2003 Cisco Systems

Published by:

Cisco Press

800 East 96th Street, 3rd Floor

Indianapolis, IN 46240 USA

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the publisher, except for the inclusion of brief quotations in a review.

Printed in the United States of America 1 2 3 4 5 6 7 8 9 0

Library of Congress Cataloging-in-Publication Number: 2002103527

ISBN: 1-58705-096-X

First Printing September 2003

## **Warning and Disclaimer**

This book is designed to provide information about Metro Ethernet. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an “as is” basis. The author, Cisco Press, and Cisco Systems, Inc., shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

## **Trademark Acknowledgments**

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc. cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

## **Feedback Information**

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members from the professional technical community.

Readers’ feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us through e-mail at [feedback@ciscopress.com](mailto:feedback@ciscopress.com). Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Publisher  
 Editor-in-Chief  
 Cisco Representative  
 Cisco Press Program Manager  
 Manager, Marketing Communications,  
 Cisco Systems  
 Cisco Marketing Program Manager  
 Production Manager  
 Development Editor  
 Copy Editor  
 Technical Editors  
 Team Coordinator  
 Cover Designer  
 Composition  
 Proofreader  
 Indexer

John Wait  
 John Kane  
 Anthony Wolfenden  
 Sonia Torres Chavez  
 Scott Miller  
 Edie Quiroz  
 Patrick Kanouse  
 Dayna Isley  
 Bill McManus  
 Mike Bernico, Mark Gallo, Giles Heron, Irwin Lazar  
 Tammi Ross  
 Louisa Adair  
 Interactive Composition Corporation  
 Gayle Johnson  
 Larry Sweazy



**Corporate Headquarters**  
 Cisco Systems, Inc.  
 170 West Tasman Drive  
 San Jose, CA 95134-1706  
 USA  
 www.cisco.com  
 Tel: 408 526-4000  
 800 553-NETS (6387)  
 Fax: 408 526-4100

**European Headquarters**  
 Cisco Systems International BV  
 Haarlerbergpark  
 Haarlerbergweg 13-19  
 1101 CH Amsterdam  
 The Netherlands  
 www-europe.cisco.com  
 Tel: 31 0 20 357 1000  
 Fax: 31 0 20 357 1100

**Americas Headquarters**  
 Cisco Systems, Inc.  
 170 West Tasman Drive  
 San Jose, CA 95134-1706  
 USA  
 www.cisco.com  
 Tel: 408 526-7660  
 Fax: 408 527-0883

**Asia Pacific Headquarters**  
 Cisco Systems, Inc.  
 Capital Tower  
 168 Robinson Road  
 #22-01 to #29-01  
 Singapore 068912  
 www.cisco.com  
 Tel: +65 6317 7777  
 Fax: +65 6317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the

**Cisco.com Web site at [www.cisco.com/go/offices](http://www.cisco.com/go/offices).**

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Czech Republic  
 Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland • Israel • Italy  
 Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal  
 Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden  
 Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Copyright © 2003 Cisco Systems, Inc. All rights reserved. CCIP, CCSP, the Cisco Arrow logo, the Cisco *Powered* Network mark, the Cisco Systems Verified logo, Cisco Unity, Follow Me Browsing, FormShare, iQ Net Readiness Scorecard, Networking Academy, and ScriptShare are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn, The Fastest Way to Increase Your Internet Quotient, and iQuick Study are service marks of Cisco Systems, Inc.; and Aironet, ASIST, BPX, Catalyst, CCDA, CCDP, CCIE, CCNA, CCNP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, the Cisco IOS logo, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Empowering the Internet Generation, Enterprise/Solver, EtherChannel, EtherSwitch, Fast Step, GigaStack, Internet Quotient, IOS, IP/TV, iQ Expertise, the iQ logo, LightStream, MGX, MICA, the Networkers logo, Network Registrar, *Packet*, PIX, Post-Routing, Pre-Routing, RateMUX, Registrar, SlideCast, SMARTnet, StrataView Plus, Stratm, SwitchProbe, TeleRouter, TransPath, and VCO are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0303R)

Printed in the USA

# About the Author

Mr. Halabi is a seasoned executive and an industry veteran with more than 18 years of experience marketing and selling to the worldwide Enterprise and Carrier networking markets. While at Cisco, he wrote the first Cisco Internet routing book, *Internet Routing Architectures*, a best-seller in the U.S. and international markets. He has held multiple executive management positions in the field of marketing, sales, and business development and has been instrumental in evolving fast-growing businesses for the Enterprise and Carrier Ethernet markets.

## About the Technical Reviewers

**Mike Bernico** is a senior networking engineer at the Illinois Century Network. In this position, he focuses primarily on network design and integrating advanced network services such as QoS, IP Multicast, IPv6, and MPLS into the network. He has also authored open-source software related to his interests in new networking technologies. He enjoys reading and spending time in the lab increasing his knowledge of the networking industry. He lives in Illinois with his wife Jayme. He can be contacted at [mike@bernico.net](mailto:mike@bernico.net).

**Mark Gallo** is a technical manager with America Online. His network certifications include Cisco CCNP and Cisco CCDP. He has led several engineering groups responsible for designing and implementing enterprise LANs and international IP networks. He has a BS in electrical engineering from the University of Pittsburgh. He resides in northern Virginia with his wife, Betsy, and son, Paul.

**Giles Heron** is the principal network architect for PacketExchange, a next-generation carrier providing Ethernet services on a global basis. He designed PacketExchange's MPLS network and has been instrumental in the development of its service portfolio. A cofounder of PacketExchange, he previously worked in the Network Architecture group at Level(3) Communications. He is coauthor of the draft-martini specification for transport of Layer 2 protocols over IP and MPLS networks and the draft-lasserre-vkompella specification for emulation of multipoint Ethernet LAN segments over MPLS, as well as various other Internet drafts.

**Irwin Lazar** is practice manager for Burton Group in its Networks and Telecom group, managing a team of consultants who advise large end-user organizations on topics including network architecture and emerging network technologies. He administers The MPLS Resource Center (<http://www.mplsrc.com>) and is the conference director for the MPLScon Conference and Exhibition. He has published numerous articles on topics relating to data networking and the Internet and is a frequent speaker on networking-related topics at many industry conferences. He holds a bachelor's degree in management information systems from Radford University and an MBA from George Mason University. He is also a Certified Information Systems Security Professional (CISSP).



# Dedications

I dedicate this book to my wonderful family, who spent many nights and weekends alone to help me finish the manuscript. To my lovely wife, Roula, I promised you after the IRA book that I wouldn't write another book. Sorry I lied. Thank you for supporting me. To my sons, Joe and Jason, I love you both for the sacrifices you had to make during the last year for me to finish this book.

# Acknowledgments

I would like to acknowledge many individuals who made this book possible. Many thanks to Giles Heron from PacketExchange for his thorough review of the material and to his many contributions to the Metro Ethernet space. I would like to thank Irwin Lazar, Mike Bernico, Mark Gallo, and Saaed Sardar for their contributions and for keeping me honest. Thanks to Andrew Malis for his initial work on this project. I also would like to thank many of the authors of the IETF RFCs and IETF drafts whose information has been used for some of the concepts and definitions in this book.

This includes the following people: Luca Martini, Nasser El-Aawar, Eric Rosen, and Giles Heron for their work on the encapsulation of Ethernet frames over IP/MPLS networks. V. Kompella, Mark Lasserre, Nick Tingle, Sunil Khandekar, Ali Sajassi, Tom Soon, Yetik Serbest, Eric Puetz, Vasile Radaoca, Rob Nath, Andrew Smith, Juha Heinanen, Nick Slabakov, J. Achirica, L. Andersson, Giles Heron, S. Khandekar, P. Lin, P. Menezes, A. Moranganti, H. Ould-Brahim, and S. Yeong-il for their work on the VPLS draft specification. K. Kompella for his original work on the DTLS draft specification. Special thanks to Daniel O. Awduche for his many contributions to traffic engineering requirements and his phenomenal work in driving multiprotocol lambda switching and GMPLS. Thanks to J. Malcolm, J. Agogbua, M. O'Dell, and J. McManus for their contributions to TE requirements. Many thanks to the CCAMP group and its many contributors to GMPLS, including Peter Ashwood Smith, Eric Mannie, Thomas D. Nadeau, Ayan Banerjee, Lyndon Ong, Debashis Basak, Dimitri Papadimitriou, Lou Berger, Dimitrios Pendarakis, Greg Bernstein, Bala Rajagopalan, Sudheer Dharanikota, Yakov Rekhter, John Drake, Debanjan Saha, Yanhe Fan, Hal Sandick, Don Fedyk, Vishal Sharma, Gert Grammel, George Swallow, Dan Guo, Kireeti Kompella, Jennifer Yates, Alan Kullberg, George R. Young, Jonathan P. Lang, John Yu, Fong Liaw, and Alex Zinin. I would also like to thank the Metro Ethernet Forum and the MPLS Forum for many of their informative references about MPLS and VPLS. I am sure I have missed many of the names of talented people who contributed indirectly to the concepts in this book, many thanks for your efforts.

Last but not least, many thanks to Cisco Systems and the Cisco Press team, John Kane, Dayna Isley, and others for supporting this project.

---

# Contents at a Glance

<b>Introduction</b>	xiii
<b>Part I</b>	<b>Ethernet: From the LAN to the MAN</b> 3
<b>Chapter 1</b>	Introduction to Data in the Metro 5
<b>Chapter 2</b>	Metro Technologies 23
<b>Chapter 3</b>	Metro Ethernet Services 45
<b>Chapter 4</b>	Hybrid L2 and L3 IP/MPLS Networks 73
<b>Part II</b>	<b>MPLS: Controlling Traffic over Your Optical Metro</b> 119
<b>Chapter 5</b>	MPLS Traffic Engineering 121
<b>Chapter 6</b>	RSVP for Traffic Engineering and Fast Reroute 133
<b>Chapter 7</b>	MPLS Controlling Optical Switches 151
<b>Chapter 8</b>	GMPLS Architecture 167
<b>Appendix</b>	SONET/SDH Basic Framing and Concatenation 193
<b>Glossary</b>	201
<b>Index</b>	211

# Table of Contents

<b>Introduction</b>	xiii	
Goals and Methods	xiii	
Who Should Read This Book?	xiv	
How This Book Is Organized	xiv	
<b>Part I</b>	<b>Ethernet: From the LAN to the MAN</b>	<b>3</b>
<b>Chapter 1</b>	<b>Introduction to Data in the Metro</b>	<b>5</b>
The Metro Network	5	
Ethernet in the Metro	8	
The Early Metro Ethernet Movers	9	
The BLECs	9	
The Metro Ethernet Carrier	10	
The Greenfield Value Proposition	11	
Bringing the Service Up in Days Rather Than Months	11	
Pay as You Grow Model	11	
Service Flexibility	11	
Lower Pricing Model	12	
The Challenges of the Greenfield Operators	12	
The U.S. Incumbent Landscape	13	
Existing Legacy TDM Infrastructure	14	
Building an All-Ethernet Data Network	14	
Pricing the Service	15	
The Incumbent Regulations	15	
The International Landscape	15	
The European Landscape	16	
The Asian Landscape	16	
A Data View of the Metro	16	
Metro Services	17	
LAN to Network Resources	18	
Ethernet L2VPN Services	19	
Ethernet Access and Frame Relay Comparison	20	
Conclusion	20	

---

**Chapter 2 Metro Technologies 23**

- Ethernet over SONET/SDH 23
  - The Role of Virtual Concatenation 25
    - Link Capacity Adjustment Scheme 27
  - EOS Used as a Transport Service 28
  - EOS with Packet Multiplexing at the Access 30
  - EOS with Packet Switching 31
    - EOS with Centralized Switching 32
    - EOS with Local Switching 32
  - EOS Interfaces in the Data Equipment 34
- Resilient Packet Ring 35
  - RPR Packet Add, Drop, and Forward 36
  - RPR Resiliency 36
  - RPR Fairness 38
- Ethernet Transport 39
  - Gigabit Ethernet Hub-and-Spoke Configuration 40
  - Gigabit Ethernet Rings 40
- Conclusion 42

**Chapter 3 Metro Ethernet Services 45**

- L2 Switching Basics 45
  - MAC Learning 46
  - Flooding 47
  - Using Broadcast and Multicast 47
  - Expanding the Network with Trunks 48
  - VLAN Tagging 49
  - The Need for the Spanning Tree Protocol 50
- Metro Ethernet Services Concepts 50
  - Ethernet Service Definition 50
  - Ethernet Service Attributes and Parameters 52
    - Ethernet Physical Interface Attribute 52
    - Traffic Parameters 52
    - Performance Parameters 54
    - Class of Service Parameters 55
    - Service Frame Delivery Attribute 56
    - VLAN Tag Support Attribute 57
    - Service Multiplexing Attribute 61
    - Bundling Attribute 61
    - Security Filters Attribute 61

- Example of an L2 Metro Ethernet Service 63
- Challenges with All-Ethernet Metro Networks 68
  - Restrictions on the Number of Customers 69
  - Service Monitoring 69
  - Scaling the L2 Backbone 69
  - Service Provisioning 69
  - Interworking with Legacy Deployments 70
- Conclusion 71

## **Chapter 4** Hybrid L2 and L3 IP/MPLS Networks 73

- Understanding VPN Components 73
- Delivering L3VPNs over IP 74
  - GRE-Based VPNs 74
  - MPLS L3VPNs 75
    - Maintaining Site Virtual Router Forwarding Tables 76
    - Using VPN-IPv4 Addresses in MPLS L3VPNs 79
    - Forwarding Traffic Across the Backbone 80
    - Applicability of MPLS L3VPNs for Metro Ethernet 80
- L2 Ethernet Services over an IP/MPLS Network 81
  - The Pseudowire Concept 83
  - PW Setup Via L2TPv3 84
  - Ethernet over MPLS—Draft-Martini 85
    - Ethernet Encapsulation 86
    - Maximum Transmit Unit 87
    - Frame Reordering 87
    - Using LDP with Directly Connected PEs 87
    - Non-Directly Connected PEs 88
  - Virtual Private LAN Service 90
    - VPLS Requirements 91
    - Signaling the VPLS Service 93
    - VPLS Encapsulation 93
    - Creating a Loop-Free Topology 93
    - MAC Address Learning 95
    - MAC Address Withdrawal 97
    - Unqualified Versus Qualified Learning 97
    - Scaling the VPLS Service Via Hierarchical VPLS 97
    - Autodiscovery 103
    - Signaling Using BGP Versus LDP 104
    - Comparison Between the Frame Relay and MPLS/BGP Approaches 105
    - L2VPN BGP Model 106

---

Example of Frame Relay Access with MPLS Edge/Core 108  
DTLS—Decoupling L2PE and PE Functionality 110

Conclusion 117

**Part II MPLS: Controlling Traffic over Your Optical Metro 119**

**Chapter 5 MPLS Traffic Engineering 121**

Advantages of Traffic Engineering 121

Pre-MPLS Traffic Engineering Techniques 123

Altering IGP Routing Metrics 123

Equal-Cost Multipath 124

Policy-Based Routing 124

Offline Design of Virtual Circuit Overlays 124

MPLS and Traffic Engineering 125

Traffic Trunks Versus LSPs 126

Capabilities of Traffic Engineering over MPLS 127

Traffic Trunk Operation and Attributes 127

Constraint-Based Routing 129

Conclusion 130

**Chapter 6 RSVP for Traffic Engineering and Fast Reroute 133**

Understanding RSVP-TE 134

RSVP LSP Tunnels 136

Label Binding and LSP Tunnel Establishment Via RSVP 137

Reservation Styles 138

Fixed Filter Reservation Style 139

Shared Explicit Reservation Style 140

Wildcard Filter Reservation Style 140

Details of the PATH Message 141

LABEL\_REQUEST Object 141

EXPLICIT\_ROUTE Object 142

RECORD\_ROUTE Object 144

SESSION\_ATTRIBUTE Object 144

FLOW\_SPEC Object 145

SENDER\_TEMPLATE Object 145

SESSION Object 145

Details of the RESV Message 145

Understanding MPLS Fast Reroute 146

End-to-End Repair 147

Local Repair 147

One-to-One Backup 148  
Facility Backup—Bypass 149

Conclusion 149

**Chapter 7** MPLS Controlling Optical Switches 151

Understanding GMPLS 151

Establishing the Need for GMPLS 152

Static and Centralized Provisioning in TDM Networks 153

The Effect of a Dynamic Provisioning Model 154

Topology and Resource Discovery 156

Path Computation and Provisioning 156

Signaling Models 157

The Overlay Model 157

The Peer Model 158

The Augmented Model 159

Label Switching in a Nonpacket World 159

Label Switching in TDM Networks 160

Signaling in a TDM Network 161

SONET/SDH LSRs and LSPs 161

The Mechanics and Function of a TDM Label 162

Label Switching in WDM Networks 163

Conclusion 164

**Chapter 8** GMPLS Architecture 167

GMPLS Interfaces 167

Modification of Routing and Signaling 168

Enhancements to Routing 168

LSP Hierarchy—Routing 170

Unnumbered Links 171

Link Bundling 172

Link Protection Types 174

Shared Risk Link Group Information 175

Interface Switching Capability Descriptor 175

Enhancements to Signaling 177

LSP Hierarchy—Signaling 177

Enhancements to Labels 180

Bandwidth Encoding 183

Bidirectional LSPs 183

Notification of Label Error 184



---

Explicit Label Control	184
Protection Information	184
Administrative Status Information	185
Separation of Control and Data Channels	185
Notify Messages	186
Inclusion of Technology-Specific Parameters	186
Link Management Protocol	187
GMPLS Protection and Restoration Mechanisms	188
Summary of Differences Between MPLS and GMPLS	189
Conclusion	191
<b>Appendix</b> SONET/SDH Basic Framing and Concatenation	193
SONET/SDH Frame Formats	193
SONET/SDH Architecture	194
SONET/SDH Concatenation	197
Contiguous Standard Concatenation	197
Virtual Concatenation	198
Conclusion	199
<b>Glossary</b>	201
<b>Index</b>	211

# Icons Used in This Book

Throughout this book, you see the following icons:



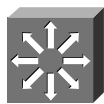
Router



Label Switch  
Router



ADM



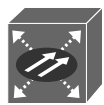
Multilayer  
Switch



Optical  
Cross-Connect



Ethernet  
Switch



Optical Transport  
Device



WDM



ISDN/Frame Relay  
Switch

---

# Introduction

Metro Ethernet—opposites attract. Ethernet is a technology that has had major success in the LAN, displacing other once-promising technologies such as Token Ring, FDDI, and ATM. Ethernet's simplicity and price/performance advantages have made it the ultimate winner, extending from the enterprise workgroup closet all the way to the enterprise backbone and data centers. The metro is the last portion of the network standing between subscribers or businesses and the vast amount of information that is available on the Internet. The metro is entrenched with legacy time-division multiplexing (TDM) and SONET/SDH technology that is designed for traditional voice and leased-line services. These legacy technologies are inadequate for handling the bandwidth demands of emerging data applications.

Ethernet in the metro can be deployed as an access interface to replace traditional T1/E1 TDM interfaces. Many data services are being deployed in the metro, including point-to-point Ethernet Line Services and multipoint-to-multipoint Ethernet LAN services or Virtual Private LAN services (VPLS) that extend the enterprise campus across geographically dispersed backbones. Ethernet can run over many metro transport technologies, including SONET/SDH, next-generation SONET/SDH, Resilient Packet Ring (RPR), and wavelength-division multiplexing (WDM), as well as over pure Ethernet transport.

Ethernet, however, was not designed for metro applications and lacks the scalability and reliability required for mass deployments. Deploying Ethernet in the metro requires the scalability and robustness features that exist only in IP and Multiprotocol Label Switching (MPLS) control planes. As such, hybrid Layer 2 (L2) and Layer 3 (L3) IP and MPLS networks have emerged as a solution that marries Ethernet's simplicity and cost effectiveness with the scale of IP and MPLS networks. With many transport technologies deployed in the metro, Ethernet services have to be provisioned and monitored over a mix of data switches and optical switches. It becomes essential to find a control plane that can span both data and optical networks. MPLS has been extended to do this task via the use of the Generalized MPLS (GMPLS) control plane, which controls both data and optical switches. Understanding these topics and more will help you master the metro space and its many intricacies.

## Goals and Methods

The goal of this book is to make you familiar with the topic of metro Ethernet—what it is, how it started, and how it has evolved. One thing is for certain: after you read this book, you will never be intimidated by the metro Ethernet topic again. You will be familiar with the different technologies, such as Ethernet switching, RPR, next-generation SONET/SDH, MPLS, and so on, in the context of metro deployments.

The industry today is divided among different pools of expertise—LAN switching, IP routing, and transport. These are three different worlds that require their own special knowledge base. LAN switching expertise is specific to individuals who come from the enterprise space, IP routing expertise is more specific to individuals who deal with public and private IP routed backbones, and transport expertise is specific to individuals who deal with TDM and optical networks. The metro blends all these areas of expertise. This book attempts to bridge the gap between enterprise LAN, IP/MPLS, and transport knowledge in the same way metro bridges the gap between enterprise networks and IP routed backbones over a blend of transport technologies.

The style of this book is narrative. It goes from simple to more challenging within each chapter and across chapters. The big picture is always presented first to give you a better view of what is being described in the chapter, and then the text goes into more details. It is possible to skip the more detailed sections of the book and still have a complete picture of the topic. I call the different levels within a chapter or across chapters “warps.” Different readers will find comfort in different warps. The main thing is to learn something new and challenging every time you enter a new warp.

## Who Should Read This Book?

The book is targeted at a wide audience, ranging from nontechnical, business-oriented individuals to very technical individuals. The different people who have interest in the subject include network operators, engineers, consultants, managers, CEOs, and venture capitalists. Enterprise directors of technology and CIOs will read the book to assess how they can build scalable virtual enterprise networks. Telecom operators will find in the book a way to move into selling next-generation data services. Engineers will augment their knowledge base in the areas of Ethernet switching, IP/MPLS, and optical networks. Salespeople will gain expertise in selling in a fast-growing metro Ethernet market. Last but not least, businesspeople will understand the topic to the level where they can make wise investments in the metro Ethernet space.

## How This Book Is Organized

This book is organized into two main parts:

- Part I—Ethernet: From the LAN to the MAN

This part of the book—Chapters 1 through 4—starts by describing the different drivers that motivated the adoption of metro Ethernet services and how they have evolved in the United States versus internationally. You will see how Ethernet has moved from the LAN into the MAN and how it is complementing existing and emerging metro technologies such as SONET/SDH, next-generation SONET, RPR, and WDM. You will then learn about the different Ethernet services, such as point-to-point Ethernet Line Services and multipoint-to-multipoint Ethernet LAN services as represented by the concept of Virtual Private LAN Service (VPLS). This part of the book explains the challenges of deploying Ethernet networks and how hybrid Ethernet and IP MPLS networks have emerged as a scalable solution for deploying L2 Ethernet VPN services.

- Part II—MPLS: Controlling Traffic over Your Optical Metro

MPLS is an important technology for scaling metro deployments. Whereas the first part of the book discusses MPLS in the context of building Layer 2 metro Ethernet VPNs, Part II—Chapters 5 through 8—explores the use of MPLS to control the traffic trajectory in the optical metro. The metro is built with data-switching, SONET/SDH, and optical-switching systems. The act of provisioning different systems and controlling traffic across packet and optical systems is difficult and constitutes a major operational expense. GMPLS has extended the use of MPLS as a universal control plane for both packet/cell and optical systems. GMPLS is one of those “warp 7” subjects. Part II first familiarizes you with the subject of traffic engineering and how the RSVP-TE signaling protocol is used to control traffic trajectory and reroute traffic in the case of failure. This makes the transition into the topic of GMPLS go smoother, with many of the basic traffic engineering in packet/cell networks already defined.



---

Chapters 1 through 8 and the appendix cover the following topics:

- **Chapter 1, “Introduction to Data in the Metro”**—The metro has always been a challenging environment for delivering data services, because it was built to handle the stringent reliability and availability needs of voice communications. The metro is evolving differently in different regions of the world, depending on many factors. For example, metro Ethernet is evolving slowly in the U.S. because of legacy TDM deployments and stiff regulations, but it is evolving quickly in other parts of the world, especially in Asia and Japan, which do not have as many legacy TDM deployments and are not as heavily regulated.
- **Chapter 2, “Metro Technologies”**—Metro Ethernet services do not necessitate an all-Ethernet Layer 2 network; rather, they can be deployed over different technologies such as next-generation SONET/SDH and IP/MPLS networks. This chapter goes into more details about the different technologies used in the metro.
- **Chapter 3, “Metro Ethernet Services”**—Ethernet over SONET, Resilient Packet Ring, and Ethernet transport are all viable methods to deploy a metro Ethernet service. However, functionality needs to be offered on top of metro equipment to deliver revenue-generating services such as Internet connectivity or VPN services. Chapter 3 starts by discussing the basics of Layer 2 Ethernet switching to familiarize you with Ethernet switching concepts. You’ll then learn about the different metro Ethernet services concepts as introduced by the Metro Ethernet Forum (MEF). Defining the right traffic and performance parameters, class of service, and service frame delivery ensures that buyers and users of the service understand what they are paying for and also helps service providers communicate their capabilities.
- **Chapter 4, “Hybrid L2 and L3 IP/MPLS Networks”**—Chapter 4 focuses first on describing a pure Layer 3 VPN implementation and its applicability to metro Ethernet. This gives you enough information to compare Layer 3 VPNs and Layer 2 VPNs relative to metro Ethernet applications. The chapter then delves into the topic of deploying L2 Ethernet services over a hybrid L2 Ethernet and an L3 IP/MPLS network. Some of the basic scalability issues that are considered include restrictions on the number of customers because of the VLAN-ID limitations, scaling the Layer 2 backbone with spanning tree, service provisioning and monitoring, and carrying VLAN information within the network.
- **Chapter 5, “MPLS Traffic Engineering”**—Previous chapters discussed how metro Ethernet Layer 2 services can be deployed over an MPLS network. Those chapters also covered the concept of pseudowires and LSP tunnels. In Chapter 5, you’ll learn about the different parameters used for traffic engineering. Traffic engineering is an important MPLS function that allows the network operator to have more control over how traffic traverses its network. This chapter details the concept of traffic engineering and its use.
- **Chapter 6, “RSVP for Traffic Engineering and Fast Reroute”**—MPLS plays a big role in delivering and scaling services in the metro, so you need to understand how it can be used to achieve traffic engineering and protection via the use of Resource Reservation Protocol traffic engineering (RSVP-TE). In this chapter, you see how MPLS, through the use of RSVP-TE, can be used to establish backup paths in the case of failure. This chapter discusses the basics of RSVP-TE and how it can be applied to establish LSPs, bandwidth allocation, and fast-reroute techniques. You’ll get a detailed explanation of the RSVP-TE messages and objects to give you a better understanding of this complex protocol.

- **Chapter 7, “MPLS Controlling Optical Switches”**—The principles upon which MPLS technology is based are generic and applicable to multiple layers of the transport network. As such, MPLS-based control of other network layers, such as the TDM and optical layers, is also possible. Chapter 7 discusses why Generalized MPLS (GMPLS) is needed to dynamically provision optical networks. You’ll learn about the benefits and drawbacks of both static centralized and dynamic decentralized provisioning models. Chapter 7 also introduces you to the different signaling models (overlay, peer, and augmented) and to how GMPLS uses labels to cross-connect the circuits for TDM and WDM networks.
- **Chapter 8, “GMPLS Architecture”**—Generalized MPLS (GMPLS) attempts to address some of the challenges that exist in optical networks by building on MPLS and extending its control parameters to handle the scalability and manageability aspects of optical networks. This chapter explains the characteristics of the GMPLS architecture, such as the extensions to routing and signaling and the technology parameters that GMPLS adds to MPLS to be able to control optical networks.
- **Appendix, “SONET/SDH Basic Framing and Concatenation”**—This appendix presents the basics of SONET/SDH framing and how the SONET/SDH technology is being adapted via the use of standard and virtual concatenation to meet the challenging needs of emerging data over SONET/SDH networks in the metro. The emergence of L2 metro services will challenge the legacy SONET/SDH network deployments and will drive the emergence of multiservice provisioning platforms that will efficiently transport Ethernet, Frame Relay, ATM, and other data services over SONET/SDH.

*This page intentionally left blank*



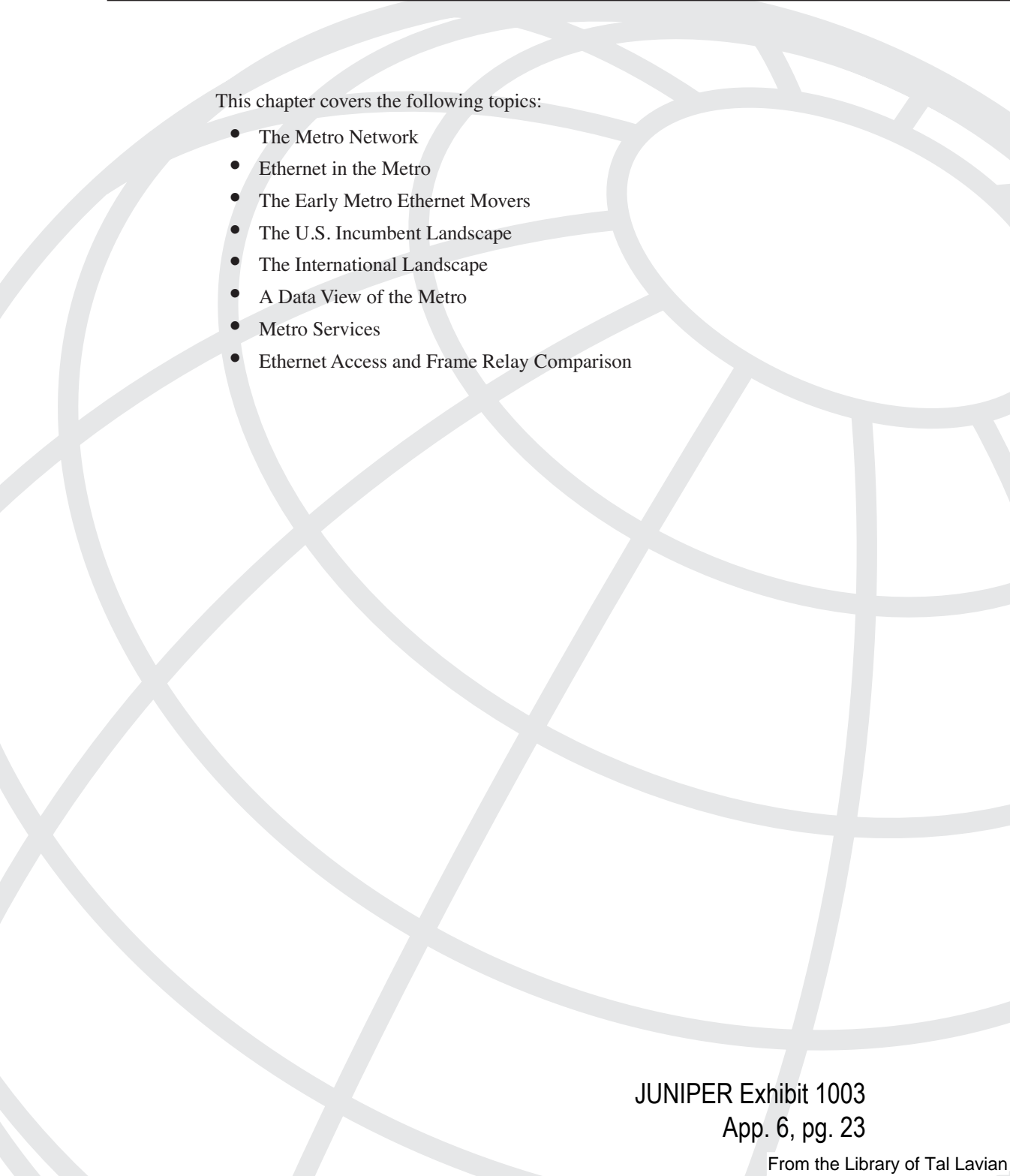

JUNIPER Exhibit 1003  
App. 6, pg. 21

From the Library of Tal Lavian



# **Ethernet: From the LAN to the MAN**

- Chapter 1 Introduction to Data in the Metro
- Chapter 2 Metro Technologies
- Chapter 3 Metro Ethernet Services
- Chapter 4 Hybrid L2 and L3 IP/MPLS Networks



This chapter covers the following topics:

- The Metro Network
- Ethernet in the Metro
- The Early Metro Ethernet Movers
- The U.S. Incumbent Landscape
- The International Landscape
- A Data View of the Metro
- Metro Services
- Ethernet Access and Frame Relay Comparison

# Introduction to Data in the Metro

---

The metro, the first span of the network that connects subscribers and businesses to the WAN, has always been a challenging environment for delivering data services because it has been built to handle the stringent reliability and availability needs of voice communications. The metro is evolving differently in different regions of the world depending on many factors, including the following:

- **Type of service provider**—Metro deployments vary with respect to the type of service providers that are building them. While regional Bell operating companies (RBOCs) are inclined to build traditional SONET/SDH metro networks, greenfield operators have the tendency to build more revolutionary rather than evolutionary networks.
- **Geography**—U.S. deployments differ from deployments in Europe, Asia Pacific, Japan, and so on. For example, while many metro deployments in the U.S. are SONET centric, China and Korea are not tied down to legacy deployments and therefore could adopt an Ethernet network faster.
- **Regulations**—Regulations tie to geography and the type of service providers. Europe, for example, has less regulation than the U.S. as far as defining the boundary between a data network and a Synchronous Digital Hierarchy (SDH) network; hence, the adoption of Ethernet over SDH deployments could move faster in Europe than in the U.S.

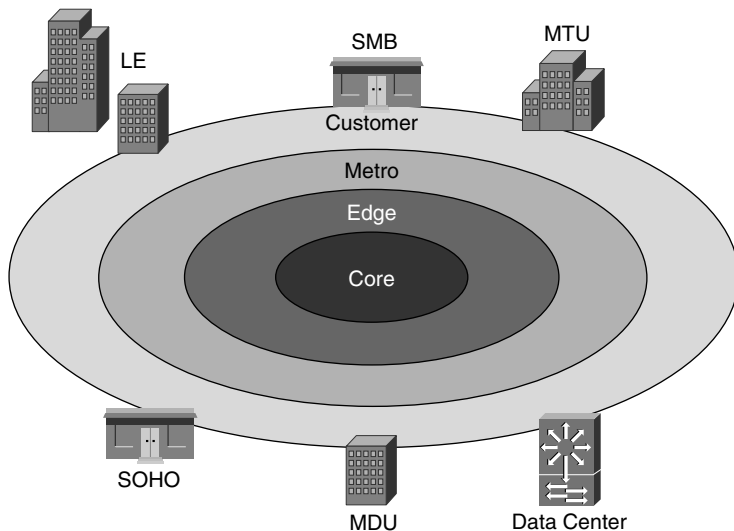
## The Metro Network

The metro is simply the first span of the network that connects subscribers and businesses to the WAN. The different entities serviced by the metro include residential and business customers, examples of which are large enterprises (LEs), small office/home office (SOHO), small and medium-sized businesses (SMBs), multitenant units (MTUs), and multidwelling units (MDUs) (see Figure 1-1).

The portion of the metro that touches the customer is called the *last mile* to indicate the last span of the carrier's network. In a world where the paying customer is at the center of the universe, the industry also calls this span the *first mile* to acknowledge that the customer comes first. An adequate term would probably be “the final frontier” because the last span

of the network is normally the most challenging and the most expensive to build and is the final barrier for accelerating the transformation of the metro into a high-speed data-centric network.

**Figure 1-1** *The Metro*



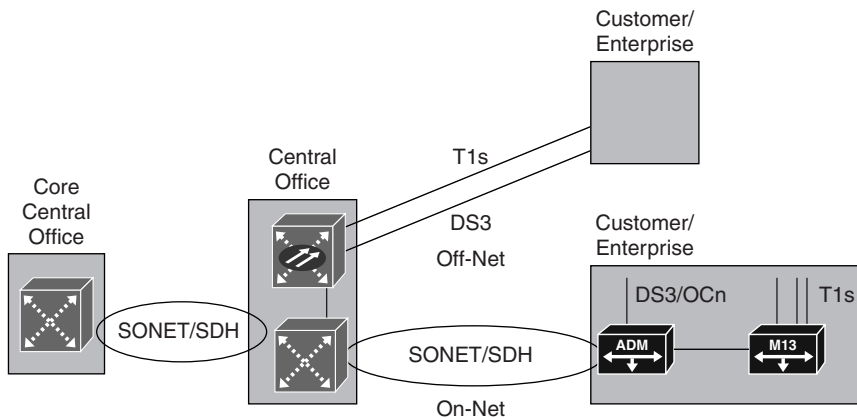
The legacy metro consists primarily of time-division multiplexing (TDM) technology, which is very optimized for delivering voice services. A typical metro network consists of TDM equipment placed in the basement of customer buildings and incumbent local exchange carrier (ILEC) central offices. The TDM equipment consists of digital multiplexers, digital access cross-connects (DACs, often referred to as digital cross-connects), SONET/SDH add/drop multiplexers (ADMs), SONET/SDH cross-connects, and so on.

Figure 1-2 shows a TDM view of a legacy metro deployment. This scenario shows connectivity to business customers for on-net and off-net networks. An *on-net* network is a network in which fiber reaches the building and the carrier installs an ADM in the basement of the building and offers T1 or DS3/OCn circuits to different customers in the building. In this case, digital multiplexers such as M13s multiplex multiple T1s to a DS3 or multiple DS3s to an OCn circuit that is carried over the SONET/SDH fiber ring to the central office (CO). In an *off-net* network, in which fiber does not reach the building, connectivity is done via copper T1 or DS3 circuits that are aggregated in the CO using DACs. The aggregated circuits are cross-connected in the CO to other core COs, where the circuits are terminated or transported across the WAN depending on the service that is being offered.

The operation and installation of a pure TDM network is very tedious and extremely expensive to deploy, because TDM itself is a very rigid technology and does not have the flexibility or the economics to scale with the needs of the customer. The cost of deploying metro networks is the sum of capital expenditure on equipment and operational expenditure. Operational expenditure includes the cost of network planning, installation, operation, and management,

maintenance and troubleshooting, and so on. What is important to realize is that these operational expenditures could reach about 70 percent of the carrier's total expenditure, which could weigh heavily on the carrier's decision regarding which products and technologies to install in the network.

**Figure 1-2** *A TDM View of the Metro*



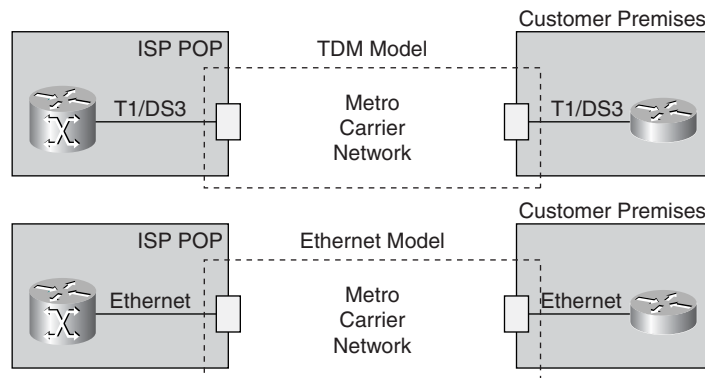
The cost of bringing up service to a customer has a huge effect on the success of delivering that service. The less the carrier has to touch the customer premises and CO equipment to deliver initial and incremental service, the higher the carrier's return on investment will be for that customer. The term *truck rolls* refers to the trucks that are dispatched to the customer premises to activate or modify a particular service. The more truck rolls required for a customer, the more money the carrier is spending on that customer.

The challenge that TDM interfaces have is that the bandwidth they offer does not grow linearly with customer demands but rather grows in step functions. A T1 interface, for example, offers 1.5 Mbps; the next step function is a DS3 interface at 45 Mbps; the next step function is an OC3 interface at 155 Mbps; and so on. So when a customer's bandwidth needs exceed the 1.5-Mbps rate, the carrier is forced to offer the customer multiple T1 (nXT1) circuits or move to a DS3 circuit and give the customer a portion of the DS3. The end effect is that the physical interface sold to the customer has changed, and the cost of the change has a major impact on both the carrier and the customer.

Moving from a T1 interface to an nXT1 or DS3/OCn requires changes to the customer premises equipment (CPE) to support the new interface and also requires changes to the CO equipment to accommodate the new deployed circuits. This will occur every time a customer requests a bandwidth change for the life of the customer connection. Services such as Channelized DS1, Channelized DS3, and Channelized OCn can offer more flexibility in deploying increments of bandwidth. However, these services come at a much higher cost for the physical interface and routers and have limited granularity. This is one of the main drivers for the proliferation of Ethernet in the metro as an access interface. A 10/100/1000 Ethernet interface scales much better from submegabit speeds all the way to gigabit, at a fraction of the cost of a TDM interface.

Figure 1-3 shows the difference between the TDM model and Ethernet model for delivering Internet connectivity. In the TDM model, the metro carrier, such as an ILEC or RBOC, offers the point-to-point T1 circuit, while the ISP manages the delivery of Internet services, which includes managing the customer IP addresses and the router connectivity in the point of presence (POP). This normally has been the preferred model for ILECs who do not want to get involved in the IP addressing and in routing the IP traffic. In some cases, the ILECs can outsource the service or manage the whole IP connection if they want to. However, this model keeps a demarcation line between the delivery of IP services and the delivery of connectivity services.

**Figure 1-3** *Connectivity: TDM Versus Ethernet*



In the Ethernet model, both network interfaces on the customer side and the ISP side are Ethernet interfaces. The ILEC manages the Layer 2 (L2) connection, while the ISP manages the IP services. From an operational perspective, this arrangement keeps the ILEC in a model similar to the T1 private-line service; however, it opens up the opportunity for the ILEC to up-sell additional service on top of the same Ethernet connection without any changes to the CPE and the network.

## Ethernet in the Metro

Ethernet technology has so far been widely accepted in enterprise deployments, and millions of Ethernet ports have already been deployed. The simplicity of this technology enables you to scale the Ethernet interface to high bandwidth while remaining cost effective. The cost of a 100-Mbps interface for enterprise workgroup L2 LAN switches will be less than \$50 in the next few years.

These costs and performance metrics and Ethernet's ease of use are motivating carrier networks to use Ethernet as an access technology. In this new model, the customer is given an Ethernet interface rather than a TDM interface.

The following is a summary of the value proposition that an Ethernet access line offers relative to TDM private lines:

- **Bandwidth scalability**—The low cost of an Ethernet access interface on both the CPE device and the carrier access equipment favors the installation of a higher-speed Ethernet interface that can last the life of the customer connection. Just compare the cost of having a single installation of a 100-Mbps Ethernet interface versus the installation of a T1 interface for 1.5-Mbps service, a T3 for 45-Mbps service, and an OC3 (155 Mbps) for 100-Mbps service. A TDM interface offering results in many CPE interface changes, many truck rolls deployed to the customer premises, and equipment that only gets more expensive with the speed of the interface.
- **Bandwidth granularity**—An Ethernet interface can be provisioned to deliver tiered bandwidth that scales to the maximum interface speed. By comparison, a rigid TDM hierarchy changes in big step functions. It is important to note that bandwidth granularity is not a function specific to Ethernet but rather is specific to any packet interface. Early deployments of metro Ethernet struggled with this function because many enterprise-class Ethernet switches did not have the capability to police the traffic and enforce SLAs.
- **Fast provisioning**—Deploying an Ethernet service results in a different operational model in which packet leased lines are provisioned instead of TDM circuit leased lines. The packet provisioning model can be done much faster than the legacy TDM model because provisioning can be done without changing network equipment and interfaces. Packet provisioning is a simple function of changing software parameters that would throttle the packets and can increase or decrease bandwidth, establish a connection in minutes, and bill for the new service.

## The Early Metro Ethernet Movers

The earliest service providers to move into the metro Ethernet space appeared in the 1999–2000 timeframe in the midst of the telecom bubble and have adopted variations of the same business model across the world.

In the U.S., the early adopters of metro Ethernet were the greenfield service providers that wanted to provide services to some niche segments, such as SMBs that are underserved by the incumbent providers. Other providers have found an opportunity in promoting cheaper bandwidth by selling Ethernet pipes to large enterprises or to other providers such as ISPs or content providers.

The greenfield operators consist of BLECs and metro operators, which are discussed next.

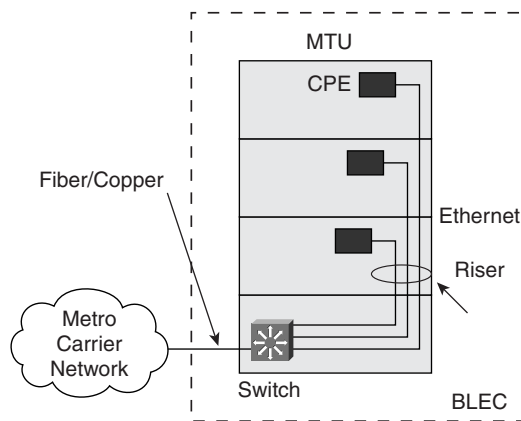
### The BLECs

The Building Local Exchange Carriers (BLECs) have adopted a retail bandwidth model that offers services to SMBs which are concentrated in large MTUs. (These are the “tall

and shiny buildings” that are usually located in concentrated downtown city areas.) The BLECs focus on wiring the inside of the MTUs for broadband by delivering Ethernet connections to individual offices. The BLECs capitalize on the fact that from the time an SMB places an order, it takes an incumbent operator three to six months to deploy a T1 circuit for that SMB. The BLECs can service the customers in weeks, days, or even hours rather than months and at much less cost.

As shown in Figure 1-4, a BLEC installs its equipment in the basement of the MTU, runs Ethernet in the risers of the building, and installs an Ethernet jack in the customer office. The customer can then get all of its data services from the Ethernet connection.

**Figure 1-4** *The BLEC Network Model*



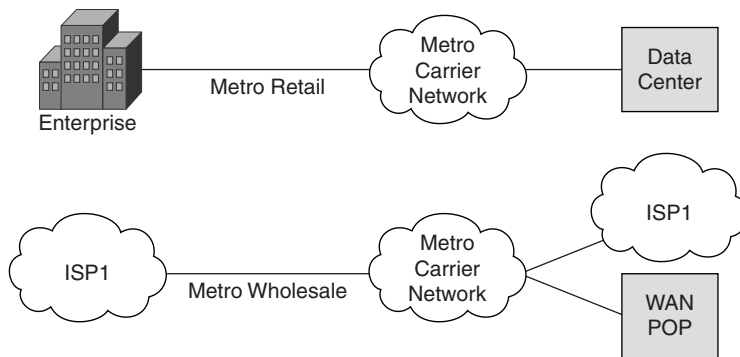
## The Metro Ethernet Carrier

Although the BLECs are considered metro operators, they specialize in servicing the MTU customers rather than building connectivity within the metro itself. The metro carriers are focused on building connectivity within the metro and then selling connectivity to BLECs, large enterprises, or even other service providers, depending on the business model. However, a lot of consolidation has occurred because metro operators have acquired BLECs, blurring the distinction between the two different providers.

Whereas some metro carriers have adopted a retail model, selling bandwidth to large enterprises, other metro carriers have adopted a wholesale model, selling bandwidth to other service providers (see Figure 1-5).

Other business plans for metro deployments target cities that want to enhance the quality of life and attract business by tying the whole city with a fiber network that connects schools, universities, businesses, financial districts, and government agencies.



**Figure 1-5** *Retail Versus Wholesale Model*

## The Greenfield Value Proposition

The following sections describe the value proposition that greenfield operators can offer to attract business away from the incumbents.

### Bringing the Service Up in Days Rather Than Months

As mentioned earlier, one of the key selling points for the metro greenfield operators is their ability to bring service up in days. However, to accomplish this, the service has to be almost ready to be brought up once the customer requests it. Greenfields spend a lot of money on idle connections, waiting for a customer to appear.

### Pay as You Grow Model

With an Ethernet connection, the customer can purchase an initial amount of bandwidth and SLA and then has the option to change the service in the future by simply calling the provider. The provider could then immediately assign the customer to a different SLA by changing the network parameters via software. Some metro operators offer their customers the ability to change their own bandwidth parameters via a web-based application.

### Service Flexibility

With an Ethernet interface, the provider can offer the customer different types of services, such as Internet access, transparent LAN service (TLS), Voice over IP (VoIP), and so on, with minimal operational overhead. Each service is provided over its own virtual LAN (VLAN) and is switched differently in the network. The different services can be sold over the same Ethernet interface or, alternatively, each service can have a separate physical interface.

## Lower Pricing Model

The initial claims for the metro Ethernet service were very aggressive. Some of the early marketing campaigns claimed “twice the bandwidth at half the price.” The quotes for 100-Mbps Ethernet connections initially ranged from \$100 per month to \$5000 per month depending on which carrier you talked to and at what time of the day you talked to them. Table 1-1 compares sample pricing for Ethernet and T1/T3 services. The Ethernet pricing might vary widely depending on the region and how aggressive the carrier gets.

**Table 1-1** *Sample Pricing Comparison for Ethernet Versus T1/T3 Private-Line Service*

<b>Greenfield</b>	<b>Incumbent</b>
1.5 Mbps at ~\$500/month	T1 (1.5 Mbps) at ~\$750/month
3 Mbps at ~\$750/month	2 * T1 at ~\$1500
45 Mbps at ~\$2250/month	T3 (45 Mbps) at ~\$6000/month

## The Challenges of the Greenfield Operators

The BLECs and metro Ethernet carriers have encountered many challenges in their business model that have hindered their success and caused a lot of them to cease to exist after the telecom downturn. This section explores several of those challenges.

### The Fight for the Building Riser

Delivering Ethernet connections to the MTU offices requires having access to the building riser, which means dealing with the building owner—although there are regulations that prevent building owners from refusing to allow access to providers. The BLECs, who normally manage to have the first access to the building, have the early field advantage in capturing real estate in the basement and the riser. Of course, how much real estate becomes available or unavailable to other BLECs who are competing for the same MTU usually depends on what percentage of the profits the building owner is receiving.

### Cost of Overbuilding the Network

Because many providers in the past operated on the “build it and they will come” theory, millions of dollars were spent on overbuilding the network, which consisted of

- Pulling fiber in the riser
- Building the last-mile connectivity
- Building the core metro network

A challenge for the BLECs is to figure out how much connectivity they need inside the building. Many BLECs have deployed as many connections as possible in the building on the hope that the BLECs will attract customers. This model has, again, resulted in a lot of money spent with no return on investment, forcing many BLECs out of business.

The premise of delivering services to customers in hours and days rather than months is made under the assumption that the BLEC has control of the network facilities inside and outside the building. The perfect solution is to have the BLEC lease or own fiber connections into the building. However, only about five percent of buildings in a metro area have access to fiber, while the rest can only be accessed via copper T1 and DS3 lines. Many BLECs are looking for the “low-hanging fruit,” buildings that are already connected via fiber. In many cases, the BLECs try to have arrangements with utility companies to pull fiber into the buildings using existing conduits. In the cases where fiber passes across the building and not into the buildings, the BLECs have to share the cost of digging up the streets with building owners or utility companies. The challenge is that the first BLEC to ask for access into a building has to share the cost of digging the trench, while the BLECs who come after can easily have access to the existing conduit.

For buildings that couldn't have fiber connectivity, the BLECs had to rely on existing copper T1 and DS3 lines to deliver bandwidth into the building. So although the BLECs were competing with the ILECs, they still had to rely on them to provide the copper lines at the ILECs' slow pace of operation.

The metro carriers that are building the metro edge and core infrastructure have sunk a lot of money into buying or leasing the fiber that connects the different points of presence. Many metro providers have locked themselves into multimillion-dollar fiber leases based on the hope that their business will grow to fill up the big pipes.

### The Breadth and Reach of Services

Metro carriers have also struggled with the different types of services that they offer and whether the service is offered on a regional or national basis. High-end customers such as large enterprises and financial institutions usually use a one-stop shop: one provider offering local and national connectivity with different types of services, such as Frame Relay or ATM VPN services. An Ethernet-only service approach with no national coverage isn't too attractive. This has forced the metro providers to remain as niche players that do not have the support and reach that the incumbents have.

### The Pricing Model

The cheap Ethernet connectivity pricing model could not be sustained. High-speed connections between 10 and 100 Mbps require a higher-speed backbone, which is expensive to build and manage. Also, the greenfield providers were still building up their customer base, and the low Ethernet pricing model did not help with a very small customer base. So Ethernet pricing for 100-Mbps connections was across the map and a trial-and-error process with prices varying by thousands of dollars depending on who you talk to.

## The U.S. Incumbent Landscape

While the greenfield operators were fast to build their metro networks, the U.S. incumbents took a sit-and-watch approach to see how the market would evolve. **JUNIPER Exhibit 1003**

Ethernet model were to succeed, it would start stealing customers from the incumbents, thereby affecting the deployment of their private-line services. Threatened by the newcomers, the RBOCs and IXC's, such as SBC, Verizon, Bellsouth, Qwest, and MCI, initiated requests for information (RFIs) to solicit information from vendors about how to deliver Ethernet services in the metro.

The challenges the incumbents face in deploying metro Ethernet are very different than the challenges of the greenfields. This section discusses some of those challenges, including the following:

- Existing legacy TDM infrastructure
- Building an all-Ethernet data network
- Pricing the services
- Regulations

## Existing Legacy TDM Infrastructure

The U.S. metro is entrenched in TDM technology, and billions of dollars have already been spent on building that network. Anyone who intends to build a new service has to consider the existing infrastructure. As inefficient as it may seem, building an Ethernet service over the legacy infrastructure might be the only viable way for some incumbents to make a first entry into the metro Ethernet business. Many of the operational models have already been built for the SONET network. Operators know how to build the network, how to manage and maintain it, and how to deliver a service and bill for it. The incumbents have the challenge of adopting their existing discipline to the metro Ethernet model.

## Building an All-Ethernet Data Network

Alternatively, some U.S. incumbents have opted (after many internal debates) to build an all-Ethernet network tailored for data services. However, as of the writing of this book, none of these networks have materialized. Incumbents, who have always dealt with SONET technology, still do not quite understand Ethernet networks. Incumbents normally build their networks and services to tailor to the masses, so any new technology they deploy needs to scale to support thousands of customers nationwide. With Ethernet's roots in enterprise networks, a big gap still exists between what the incumbents need and what existing Ethernet switches, or existing Ethernet standards, have to offer. Incumbents are also unfamiliar with how to manage an Ethernet network, price the service, and bill for it. All of these factors have contributed to the delay in the deployment of such networks.

The deficiencies in Ethernet technology and Ethernet standards in dealing with the metro scalability and availability requirements were one of the main reasons for the proliferation of MPLS in the metro. This topic will be explained in more detail in Chapter 4, "Hybrid L2 and L3 IP/MPLS Networks."

## Pricing the Service

For the incumbents, pricing the metro Ethernet services is an extremely challenging exercise. Incumbents that are selling T1 and DS3 connectivity services would be competing with themselves by offering Ethernet services. A very aggressive Ethernet pricing model would jeopardize the sales of T1 and DS3 lines and disrupt the incumbent's business model.

For incumbents selling T1/E1 and DS3 services, their Ethernet pricing model has to do the following to succeed:

- Move hand in hand with existing pricing for legacy services to avoid undercutting the legacy services.
- Offer different levels of services with different price points, in addition to the basic connectivity service. Metro Ethernet services present a good value proposition for both the customer and carrier. The customer can enjoy enhanced data services at higher performance levels, and the carrier can benefit from selling services that it otherwise wouldn't have been able to sell with a simple TDM connection. So the carrier can actually sell the Ethernet connection at a lower price than the legacy connection, based on the hope that the additional services will eventually result in a more profitable service than the legacy services.

## The Incumbent Regulations

Another area that challenges the deployment of metro Ethernet services in the U.S. are the regulations that the incumbent carriers have and the delineation between the regulated and unregulated operation inside the same carriers. The regulated portions of the incumbents deal mainly with transport equipment and have rules and guidelines about the use and the location of data switching equipment. The unregulated portion of the incumbent normally has enough flexibility to deploy a mix of hybrid data switching and transport equipment without many ties.

These regulations have created a big barrier inside the incumbents and have created two different operational entities to deal with data and transport. The deployment of new data services such as metro Ethernet will prove to be challenging in the U.S. because such services require a lot of coordination between the data operation and the transport operation of the same incumbent carrier.

## The International Landscape

In 2000, while the U.S. market was bubbling with greenfield operators building their metro networks and challenging the almighty RBOCs and IXCs, the metro Ethernet market was taking its own form and shape across the globe. What was different about the rest of the world was the lack of venture capital funding that had allowed new greenfield providers to mushroom in the U.S.

## The European Landscape

In Europe, the first activities in metro Ethernet occurred in Scandinavia, specifically Sweden. Telia, the largest Swedish telecom provider, had submitted an RFI for metro Ethernet services. Unlike the U.S., where the providers were focusing on T1 private-line replacement, the target application in Sweden was residential. Many MDU apartment complexes were located in concentrated residential areas, and many of the new developments had fiber already deployed in the basements of the MDUs. Ethernet services seemed like the perfect vehicle to deliver value-added services such as converged voice, data, and video applications. A single Ethernet connection to an MDU could provide Internet access, VoIP, video on demand, and so on.

Also across Europe, a handful of greenfield operators had very aggressive plans to deploy metro Ethernet services, but most faced the same challenges as the U.S. greenfield operators. In pockets of Europe such as Italy, large players such as Telecom Italia were experimenting with an all-Ethernet metro for residential customers.

In general, however, the European metro is entrenched in SDH technology and, like the U.S., has invested in legacy TDM deployments. This puts the big European providers in the same challenging position as the U.S. incumbents in dealing with service cannibalization and the cost of a new buildout. However, Europe differs from the U.S. in that it doesn't have stringent regulations that require a strict boundary between the operation of data switching equipment and SDH transport equipment, which could play a big role in the shift toward metro Ethernet buildouts.

## The Asian Landscape

The metro Ethernet landscape in Asia is very different than in the U.S. and Europe. Japan, Korea, and China will prove to be the major players in the deployment of all-Ethernet metro services. One of the major reasons is that these countries haven't invested as much in SONET or SDH and, thus, have a cleaner slate than the U.S. and Europe from which to deploy new data services in the metro.

Many metro Ethernet deployments have already been announced and deployed by big telecom providers such as Korea Telecom SK and others. China will also emerge as a big player in this market after the restructuring of China Telecom into different entities, China Netcom, Unicom, and Railcom.

In Japan, tough competition between telecom providers has driven the cost of private-line services lower than in most other countries. Japan is also a leader in all-metro Ethernet deployments for multimedia services.

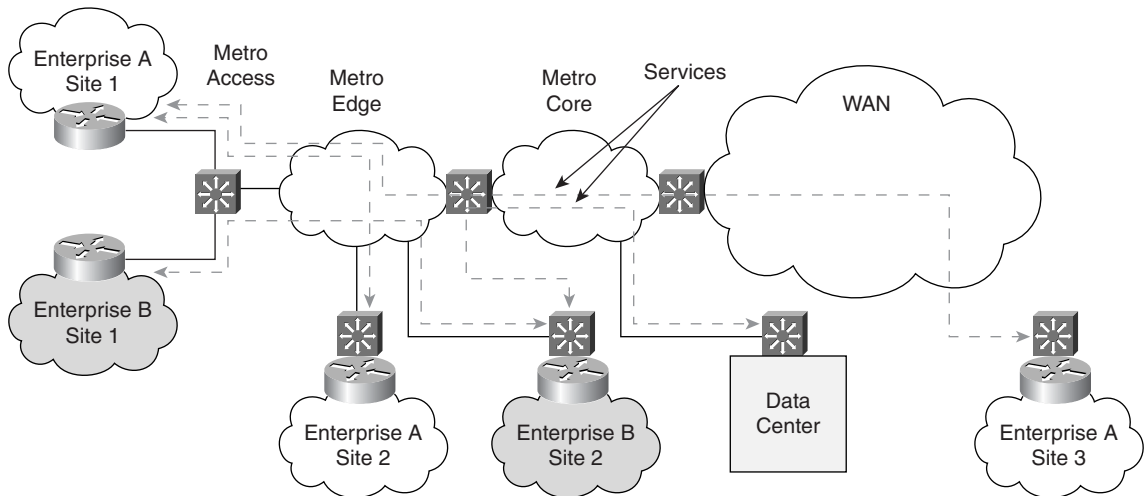
## A Data View of the Metro

A data view of the metro puts in perspective the different metro services and how they are offered by the different providers.

Figure 1-6 shows a view of the metro with the emphasis on the data access, data aggregation, and service delivery. As you can see, the metro is divided into three segments:

- **Metro access**—This segment constitutes the last-mile portion, which is the part of the network that touches the end customer. For business applications, for example, access equipment resides in a closet in the basement of the enterprise or MTU.
- **Metro edge**—This segment constitutes the first level of metro aggregation. The connections leaving the buildings are aggregated at this CO location into bigger pipes that in turn get transported within the metro or across the WAN.
- **Metro core**—This segment constitutes a second level of aggregation where many edge COs are aggregated into a core CO. In turn, the core COs are connected to one another to form a metro core from which traffic is overhauled across the WAN.

**Figure 1-6** *Data View of the Metro*



The terminology and many variations of the metro can be confusing. In some cases, there is only one level of aggregation; hence, the building connections are aggregated into one place and then directly connected to a core router. In other scenarios, the metro core CO, sometimes called the metro hub, co-locates with the wide-area POP.

## Metro Services

The metro services vary depending on the target market—commercial or residential—and whether it is a retail service or a wholesale service. The following list gives a summary of some of the metro services that are promoted:

- Internet connectivity
- Transparent LAN service (point-to-point LAN to LAN)

- L2VPN (point-to-point or multipoint-to-multipoint LAN to LAN)
- LAN to network resources (remote data center)
- Extranet
- LAN to Frame Relay/ATM VPN
- Storage area networks (SANs)
- Metro transport (backhaul)
- VoIP

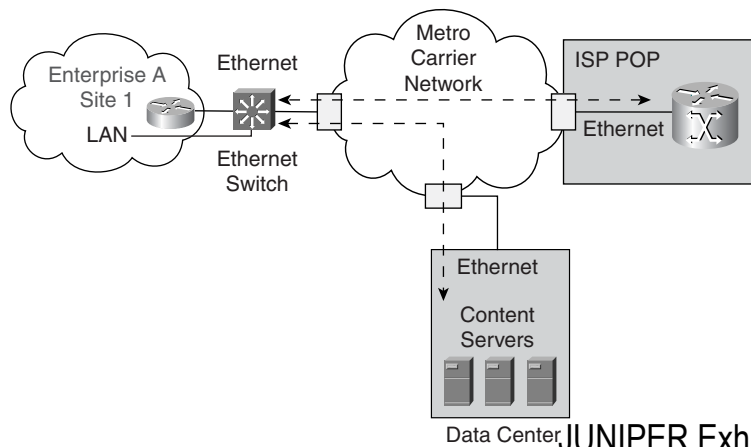
Some of these services, such as Internet connectivity and TLS, have been offered for many years. The difference now is that these services are provided with Ethernet connectivity, and the carriers are moving toward a model in which all of these services can be offered on the same infrastructure and can be sold to the same customer without any major operational overhead. This introduces an excellent value proposition to both the customer and the carrier. The services are provisioned through transporting the application over point-to-point or multipoint-to-multipoint L2 connections. The following sections discuss some of these services in greater detail.

## LAN to Network Resources

Earlier, in the section “The Metro Network,” you saw how Internet service can be delivered by installing at the customer premises an Ethernet connection rather than a T1 TDM connection. After the Ethernet connection is installed at the end customer, the ILEC can sell different services to the customer, such as LAN to network resources. An example of such a service is one that enables an enterprise to back up its data in a remote and secure location for disaster recovery.

Figure 1-7 shows that in addition to Internet service, the customer can have a data backup and disaster recovery service that constantly backs up data across the metro.

**Figure 1-7** LAN to Network Resources





For new data networks in which the connectivity is done via gigabit and 10 gigabit pipes, the metro can be transformed into a high-speed LAN that offers bandwidth-intensive applications that would not normally be feasible to deploy over legacy TDM infrastructure.

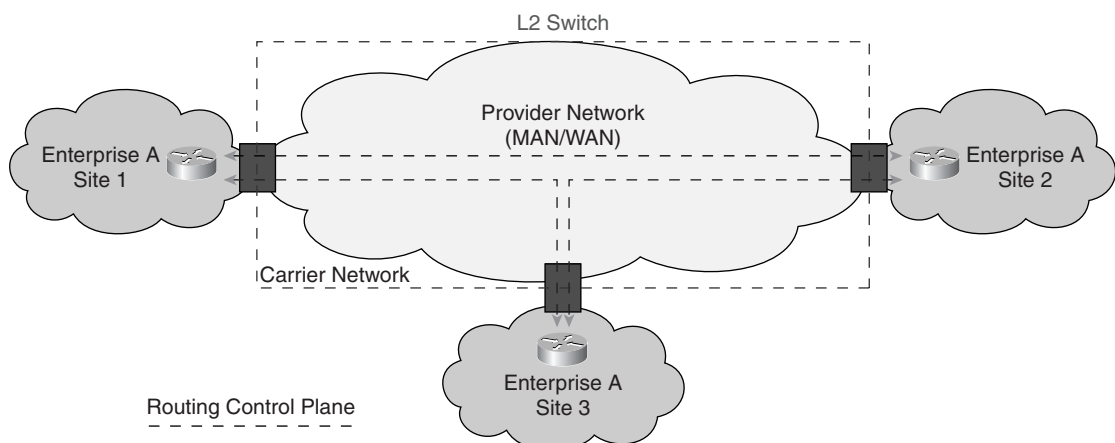
As previously mentioned, the service in the metro will take many shapes and forms depending on the target customer. The same LAN to network resources model could be applied toward residential applications, enabling the ILECs to start competing with cable companies in distributing multimedia services. In a residential application, video servers would be located in a metro POP and residential MDU customers could access high-speed digital video on demand over an Ethernet connection. While these services still seem futuristic in the U.S., the international landscape soon could be very different in Europe (particularly Sweden), Japan, and Korea, where the fast deployment of Ethernet networks is already making these applications a reality.

## Ethernet L2VPN Services

You may have noticed that many of the services mentioned are pure L2 services that offer connectivity only. This is similar to legacy Frame Relay and ATM services, where the Frame Relay/ATM connection offers a pure L2 pipe and the IP routed services can ride on top of that pipe.

Figure 1-8 shows a carrier deploying an Ethernet L2VPN service. The carrier network behaves as an L2 Ethernet switch that offers multipoint-to-multipoint connectivity between the different customer sites. The customer can benefit from running its own control plane transparently over the carrier's network. The customer routers at the edge of the enterprise could exchange routing protocols without interference with the carrier routing, and the carrier would not have to support any of the customer's IP addressing. An important observation is that while the carrier's network behaves like an L2 Ethernet switch, the underlying technology and the different control planes used in the carrier network are not necessarily based on Ethernet or a Layer 2 control plane.

**Figure 1-8** L2VPN services



## Ethernet Access and Frame Relay Comparison

Frame Relay VPN services have been widely accepted and have proven to be very cost effective compared to point-to-point private-line service. In essence, Ethernet services can be considered the next-generation of Frame Relay because they provide most of the benefits of Frame Relay with better scalability as far as providing higher bandwidth and multipoint-to-multipoint connectivity services. The following list shows some of the similarities and dissimilarities between an Ethernet and a Frame Relay service:



- **Interface speed**—Frame Relay interface speeds range from sub-T1 rates up to OCn speeds. However, Frame Relay has been widely deployed at the lower sub-T1, T1, and DS3 speeds. An Ethernet interface can run at up to 10 Gbps.
- **Last-mile connectivity**—Ethernet services will find better acceptance in on-net deployments (where fiber reaches the building), irrespective of the transport method (as will be explained in the next chapter). Frame Relay has the advantage of being deployed in off-net applications over existing copper T1 and DS3 lines, which so far constitutes a very high percentage of deployments. There are existing efforts in forums, such as the Ethernet in the First Mile (EFM) forum, to run Ethernet directly over existing copper lines. It is unknown at this point whether such a deployment would find acceptance compared to a traditional Frame Relay service.
- **Virtual circuit support**—Both Ethernet and Frame Relay offer a multiplexed interface that allows one customer location to talk to different locations over the same physical interface. The VLAN notion of Ethernet is similar to the Frame Relay permanent virtual circuit (PVC).
- **Multipoint connectivity**—An obvious difference between Frame Relay and Ethernet is that Frame Relay virtual circuits are point-to-point circuits. Any point-to-multipoint or multipoint-to-multipoint connectivity between sites is done via the provisioning of multiple point-to-point PVCs and routing between these PVCs at a higher layer, the IP layer. With Ethernet, the VLAN constitutes a broadcast domain, and many sites can share multipoint-to-multipoint connectivity at L2.
- **L2 interface**—A very important benefit that both Frame Relay and Ethernet offer is the ability to keep the separation between the network connectivity at L2 and the higher-level IP application, including L3 routing. This allows the customer to have control over its existing L2 or L3 network and keep a demarcation between the customer's network and the carrier's network.

## Conclusion

The proliferation of data services in the metro is already taking place. You have seen in this chapter how metro data services and specifically Ethernet services are making their move into the metro. The greenfield metro operators have had quite an influence on this shift by putting pressure on traditional metro operators, such as the ILECs. While metro Ethernet is evolving

slowly in the U.S. due to legacy TDM deployments and regulations, it has found good success in different parts of the world, especially in Asia and Japan. Metro Ethernet services offer an excellent value proposition both to service providers and to businesses and consumers. Metro Ethernet services will reduce the recurring cost of service deployment while offering much flexibility in offering value-added data services.

Metro Ethernet services do not necessitate an all-Ethernet L2 network; rather, they can be deployed over different technologies such as next-generation SONET/SDH and IP/MPLS networks. Chapter 2, "Metro Technologies," goes into more details about the different technologies used in the metro.



This chapter covers the following topics:

- Ethernet over SONET/SDH (EOS)
- Resilient Packet Ring (RPR)
- Ethernet Transport

# Metro Technologies

---

Metro Ethernet services and applications do not necessarily require Ethernet as the underlying transport technology. The metro can be built on different technologies, such as

- Ethernet over SONET/SDH (EOS)
- Resilient Packet Ring (RPR)
- Ethernet Transport

## Ethernet over SONET/SDH

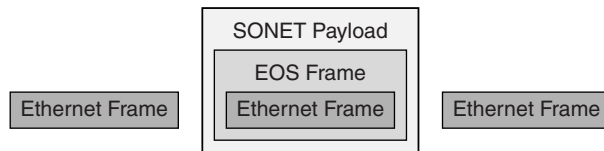
Many incumbent carriers in the U.S. and Europe have already spent billions of dollars building SONET/SDH metro infrastructures. These carriers would like to leverage the existing infrastructure to deliver next-generation Ethernet services. For such deployments, bandwidth management on the network is essential, because of the low capacity of existing SONET/SDH rings and the fact that they can be easily oversubscribed when used for data services.

Incumbents who want to deploy EOS services face tough challenges. Traditionally, for RBOCs and ILECs in the U.S., there is a clear-cut delineation between transport and data. The regulated part of the organization deals with transport-only equipment, not data equipment. With EOS, the equipment vendors blur the line between data and transport, which creates a problem for the adoption of the new technology. So, it is worth spending some time explaining the EOS technology itself.

The benefit of EOS is that it introduces an Ethernet service while preserving all the attributes of the SONET infrastructure, such as SONET fast restoration, link-quality monitoring, and the use of existing SONET OAM&P network management. With EOS, the full Ethernet frame is still preserved and gets encapsulated inside the SONET payload at the network ingress and gets removed at the egress.

As Figure 2-1 shows, the entire Ethernet frame is encapsulated inside an EOS header by the EOS function of the end system at the ingress. The Ethernet frame is then mapped onto the SONET/SDH Synchronous Payload Envelope (SPE) and is transported over the SONET/SDH ring. The Ethernet frame is then extracted at the EOS function on the egress side.

**Figure 2-1** *Ethernet over SONET*



There are two standardized ways to transport Ethernet frames over a SONET/SDH network:

- **LAPS**—Ethernet over the Link Access Procedure SDH is defined by the ITU-T, which published the X.86 standard in February 2001. LAPS is a connectionless protocol similar to High-Level Data Link Control (HDLC).
- **GFP**—Generic Framing Procedure is also an ITU standard that uses the Simple Data Link (SDL) protocol as a starting point. One of the differences between GFP and LAPS is that GFP can accommodate frame formats other than Ethernet, such as PPP, Fiber Channel, fiber connectivity (FICON), and Enterprise Systems Connection (ESCON).

The EOS function can reside inside the SONET/SDH equipment or inside the packet switch. This creates some interesting competing scenarios between switch vendors and transport vendors to offer the Ethernet connection.

Figures 2-2, 2-3, and 2-4 show different scenarios for the EOS connection. In Figure 2-2, the EOS function is inside the ADM. This is normally done via a combination framer/mapper that supports EOS and is placed on a line card or daughter card inside the ADM. The EOS mapping function adds an X.86 or GFP wrapper around the whole Ethernet frame, and the framing function encapsulates the frame in the SONET/SDH SPE. From then on, the SONET/SDH SPE is transported across the SONET/SDH ring and gets peeled off on the egress side. ADMs that contain the EOS function plus other functions such as virtual concatenation (discussed in the next section) are called next-generation ADMs. Figure 2-3 places the EOS function inside the switch.

**Figure 2-2** *EOS Function Inside the ADM*



The difference here is that the data equipment and the transport equipment are two different entities that can be owned by different operational groups within the same carrier. This makes it much easier for regulated and unregulated entities within the carrier to deploy a new service. The regulated group's sole responsibility is to provision SONET/SDH circuits, as they would do for traditional voice or leased-line circuits. The unregulated group in turn deploys the

higher-layer data services. It is also worth mentioning that in this scenario, the Ethernet switch that delivers the data services has full control of the SONET/SDH tributaries. This is in contrast to Figure 2-2, in which the SONET/SDH tributaries are terminated inside the ADM, and the Ethernet switch sees only a concatenated Ethernet pipe. Figure 2-4 combines the packet-switching, ADM, and EOS functions in the same equipment.

**Figure 2-3** *EOS Function Inside the Switch*



For equipment efficiency, this is the optimal solution; however, the deployment of such systems can be challenging if strict operational delineation between packet and transport exists. Such deployments are occurring in the U.S. by smaller competitive telecom providers and by the unregulated portion of the RBOCs/ILECs that do not have many restrictions about data versus transport. Deployments of such systems in Europe are more prevalent because Europe has fewer restrictions than the U.S.

**Figure 2-4** *EOS and Switching Functions Inside the ADM*



EOS introduces some bandwidth inefficiencies in deploying metro Ethernet services because of the coarse bandwidth granularity of SONET/SDH circuits and the bandwidth mismatch with the sizes of Ethernet pipes. Virtual concatenation (VCAT) is a mechanism used to alleviate such inefficiencies, as discussed next.

## The Role of Virtual Concatenation

Virtual concatenation is a measure for reducing the TDM bandwidth inefficiencies on SONET/SDH rings. With standard SONET/SDH concatenation, SONET/SDH pipes are provisioned with coarse granularity that cannot be tailored to the actual bandwidth requirement. The TDM circuits are either too small or too large to accommodate the required bandwidth. On a SONET/SDH ring, once the circuit is allocated, the ring loses that amount of bandwidth whether the bandwidth is used or not.

Appendix A, “SONET/SDH Basic Framing and Concatenation,” briefly describes SONET/SDH and the different terminology you see throughout this chapter.

With VCAT, a number of smaller pipes are concatenated and assembled to create a bigger pipe that carries more data per second. Virtual concatenation is done on the SONET/SDH layer (L1)

itself, meaning that the different individual circuits are bonded and presented to the upper network layer as one physical pipe. Virtual concatenation allows the grouping of  $n * \text{STS}/\text{STM}$  or  $n * \text{VT}/\text{VC}$ , allowing the creation of pipes that can be sized to the bandwidth that is needed.

Figure 2-5 highlights the bandwidth efficiency that VCAT can provide. If standard concatenation is used and the bandwidth requirement is for 300 Mbps (about six STS-1s), the carrier has the option of provisioning multiple DS3 interfaces and using packet multiplexing techniques at the customer premises equipment (CPE) to distribute the traffic over the interfaces. (Note that a DS3 interface is the physical interface that runs at a 45-Mbps rate, while an STS-1 is a SONET envelope that can carry 50 Mbps.) Provisioning multiple DS3s at the CPE is normally inefficient, because it increases the cost, does not guarantee the full bandwidth (because of packet load-sharing techniques), and restricts the packet flow to 45 Mbps (because the individual physical circuits are restricted to DS3 bandwidth). The other alternative is for the carrier to allocate a full OC12 (12 STS-1s); this causes the carrier to lose revenue from selling six STS-1s, because they are allocated to a particular customer and cannot be used for other customers on the ring. With virtual concatenation, the carrier can provision a 300 Mbps pipe by bonding six STS-1s as one big pipe—hence no wasted bandwidth.

**Figure 2-5** *Virtual Concatenation*

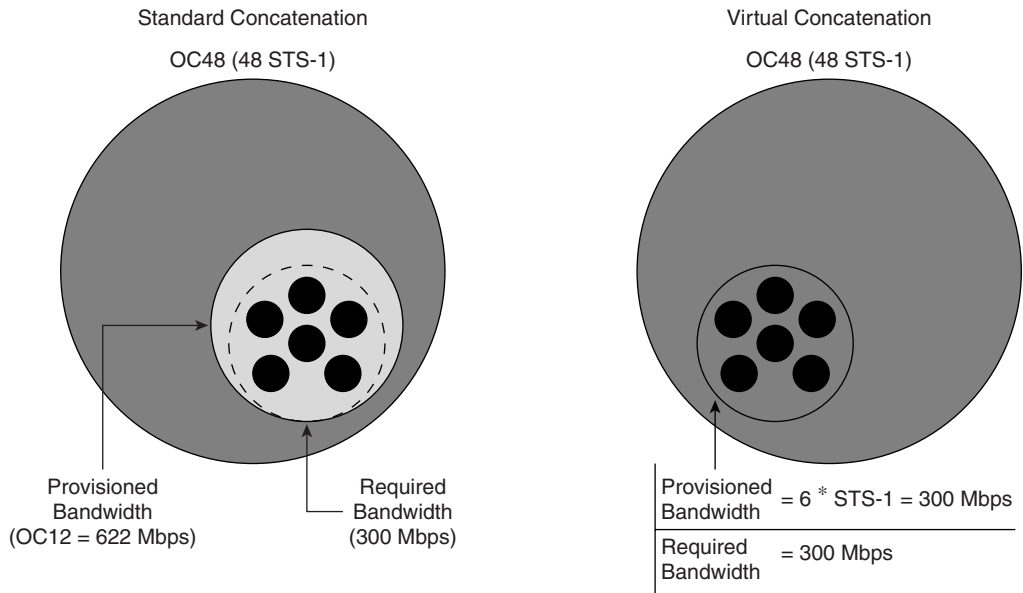
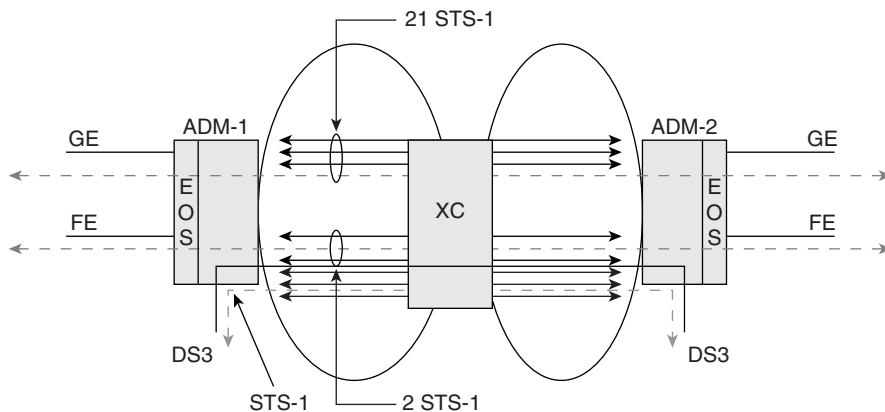


Figure 2-6 shows an example of how multiple services such as Ethernet connectivity services and traditional TDM services can be carried over the same SONET/SDH infrastructure. If the SONET/SDH equipment supports VCAT, a Gigabit Ethernet interface can be carried over a concatenated 21 STS-1 pipe, another Fast Ethernet (FE) 100-Mbps interface can be carried over two STS-1s, and a traditional DS3 interface can be carried over a single STS-1. In many cases, the speed of the Ethernet interface does not have to match the speed on the SONET/SDH side.



A Fast Ethernet 100-Mbps interface, for example, can be carried over an STS-1 (50 Mbps), two STS-1s, or three STS-1s. To handle this oversubscription, throttling of data and queuing of packets or some kind of data backoff need to happen to minimize packet loss.

**Figure 2-6** *Transporting Ethernet over SONET*



Most rings today support channelization down to the STS-1 (DS3) level and can cross-connect circuits at that level. For T1 services, M13 multiplexers are used to aggregate multiple T1 lines to a DS3 before transporting them on the ring. SONET/SDH equipment that operates at the VT/VC level is starting to be deployed by some RBOCs, which means that with virtual concatenation, circuits of  $n * VT/VC$  size can be provisioned.

The EOS and VCAT functions are implemented at the entry and exit points of the SONET/SDH infrastructure, and not necessarily at every SONET/SDH station along the way. In Figure 2-6, ADMs 1 and 2 support the EOS and VCAT functions, while the cross-connect (XC) that connects the two rings functions as a traditional cross-connect. However, for VCAT to be effective, the SONET/SDH equipment on the ring has to be able to cross-connect the tributaries supported by the VCAT; otherwise, the bandwidth savings on the ring are not realized. So, if the equipment on the ring supports the allocation of STS-1 circuits and higher, the smallest circuit that can be allocated is an STS-1 circuit. If the terminating equipment supports VCAT to the VT 1.5 level (T1), a full STS-1 bandwidth is still wasted on the ring even if the CPE is allocated  $n * VT 1.5$  via VCAT. In Figure 2-6, for example, if ADMs 1 and 2 support VCAT down to the VT 1.5 (T1) level, and the cross-connect can cross-connect only at the STS-1 (DS3) level, the savings are not realized.

## Link Capacity Adjustment Scheme

Virtual concatenation is a powerful tool for efficiently grouping the bandwidth and creating pipes that match the required bandwidth. However, the customer bandwidth requirement could change over time, which requires the SONET/SDH pipes to be resized. This could cause

network disruption as more SONET/SDH channels are added or removed. Link Capacity Adjustment Scheme (LCAS) is a protocol that allows the channels to be resized at any time without disrupting the traffic or the link. LCAS also performs connectivity checks to allow failed links to be removed and new links to be added dynamically without network disruption.

The combination of EOS, VCAT, and LCAS provides maximum efficiency when deploying Ethernet services over SONET.

---

**NOTE**

Virtual concatenation and EOS are orthogonal technologies, meaning that they are totally independent. EOS is a mapping technology that runs over standard concatenation and VCAT; however, the full benefits are achieved if done over the latter.

---

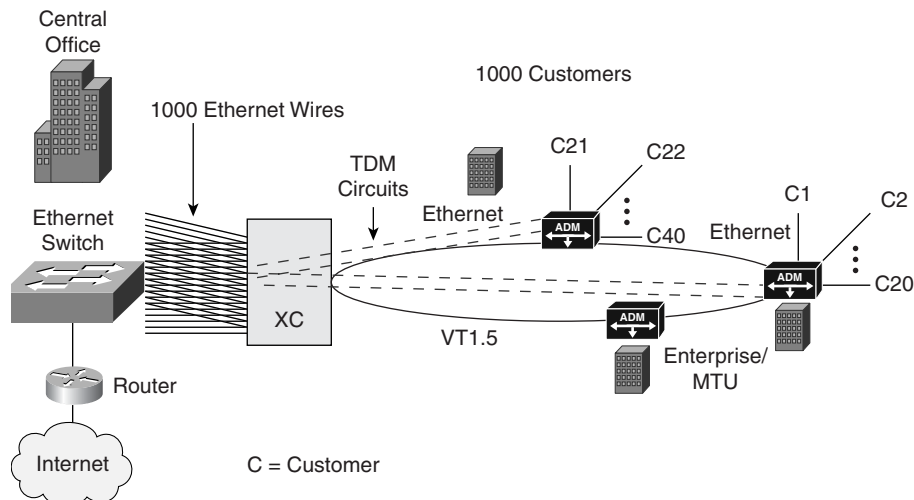
The following sections describe different scenarios in which EOS is used as a pure transport service or is applied in conjunction with packet switching.

## **EOS Used as a Transport Service**

Ethernet over SONET/SDH by itself is still a transport service with an Ethernet interface, similar to the traditional private-line service with a T1, DS3, or OCn interface. EOS offers what is comparable to a point-to-point packet leased-line service. It provides an easy migration for carriers that sell transport to get their feet wet with Ethernet services. EOS is a “packet mapping” technology, not a “packet switching” technology, and does not offer the packet multiplexing that is needed for the aggregation and deaggregation of services. To deliver enhanced switched data services, you need to introduce packet-switching functionality into the metro equipment.

The lack of packet multiplexing for the EOS service and the fact that thousands of point-to-point circuits need to be provisioned between the customers and the central office (CO) create a problem in the aggregation of services in large-scale deployments. Each individual EOS circuit could be presented as a separate Ethernet interface in the CO. With thousands of customers getting an EOS circuit, the CO would have to terminate thousands of individual Ethernet wires. Imagine if the Public Switched Telephone Network (PSTN) were still operating with each customer line terminated inside the CO as a physical wire rather than as a logical circuit. This would create a big patch-panel effect and a nightmare for provisioning and switching between circuits. This patch-panel effect presents a scalability limitation for large-scale EOS deployments, as explained next.

Figure 2-7 shows a scenario in which a carrier is using EOS to sell a basic Internet-connectivity service. A SONET/SDH metro access ring connects multiple enterprise and multitenant unit (MTU) buildings to a CO location. Next-generation ADMs in the basements of the buildings provide 100-Mbps Ethernet connections that connect to the individual routers at each customer premise. The ring itself, in this example, allows channelization down to the VT 1.5 (T1) level, and each Ethernet connection is mapped to one or  $n * VT 1.5$  circuits.

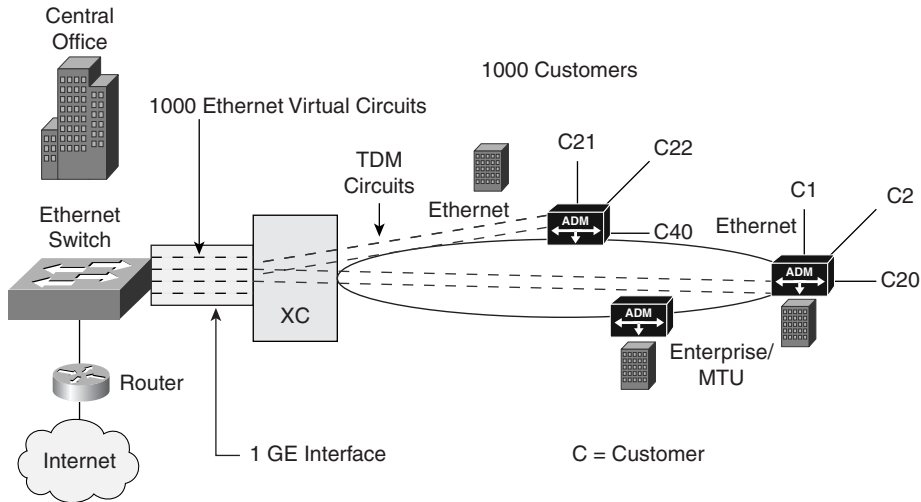
**Figure 2-7** EOS Inside Transport Equipment

For every customer who is provisioned with an Ethernet interface, an Ethernet interface is extended out of the XC at the CO, because the XC works at the TDM level, and each circuit has to be terminated individually. The individual Ethernet interfaces are then connected to an Ethernet switch that aggregates the traffic toward the ISP router. This means that if a building has 20 customers, 20 different circuits have to be provisioned for that building and have to be terminated in the CO. If the CO supports 50 buildings with 20 customers per building, 1000 TDM circuits have to be provisioned, and hence 1000 Ethernet interfaces have to be terminated in the CO. This model is very inefficient and does not scale well in terms of equipment or management. The XC will be loaded with physical Fast Ethernet interfaces, and the physical connectivity is unmanageable. The logical solution for this problem is to introduce aggregation techniques inside the cross-connect using Ethernet VLANs and to aggregate multiple Ethernet circuits over a single Gigabit or 10 Gigabit Ethernet interface where each circuit is individually identified. While such techniques are possible, they would mean more involvement of transport vendors on the data side, which is challenging from an operational perspective, especially in the U.S.

Figure 2-8 shows an example in which the XC aggregates the different EOS circuits over a single Ethernet interface that connects to an Ethernet switch. For this to happen, the XC needs to be able to logically separate the individual EOS circuits when presenting them to the Ethernet switch. This needs to be done because the traffic sent from the Ethernet switch to the XC over the GE port needs to be tagged with the right circuit ID to reach the correct destination. One method is to have the XC tag individual circuits with a VLAN ID before sending the traffic to the Ethernet switch. Other current implementations put the whole Ethernet bridging function inside the XC itself to allow the multiple EOS streams to be aggregated over a single interface when leaving the transport equipment. This has all the signs of fueling an ongoing industry

debate over what functions reside in which equipment as the data vendors and transport vendors start stepping on each other's toes.

**Figure 2-8** EOS Aggregation Inside Transport Equipment

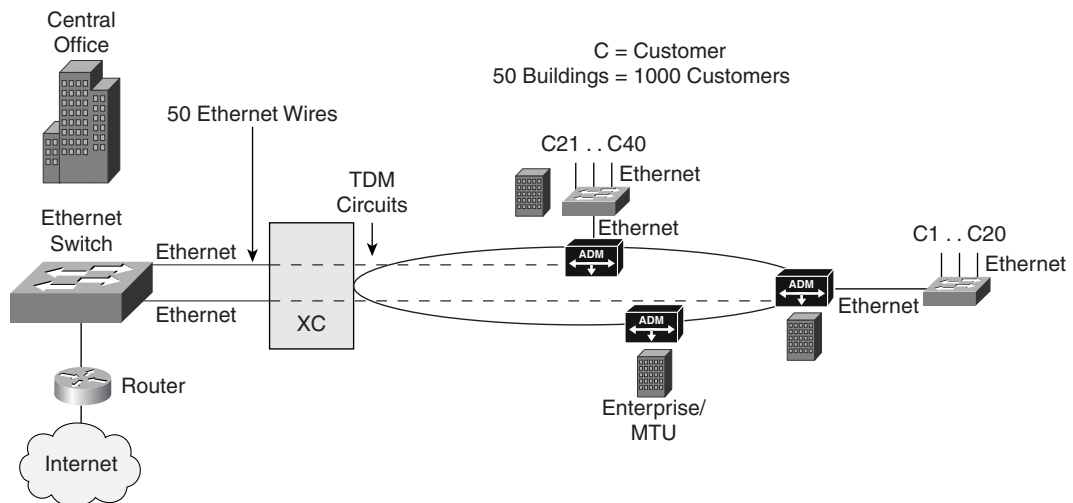


An obvious benefit of a transport service that gives each customer its own TDM circuit is that the customer is guaranteed the full bandwidth. The metro carriers that sell SONET/SDH circuits have dealt with this model for years and know full well how to substantiate the SLAs they sell. When this model is used with VCAT, which enables carriers to tailor the size of the circuit to the customer's need, carriers can realize great bandwidth efficiency and offer firm QoS guarantees. However, you need to weigh this with the complexity of managing the multitude of additional TDM circuits that have to be provisioned, because all these new circuits need to be cross-connected in the network.

### **EOS with Packet Multiplexing at the Access**

The previous example assumes that each customer in the building gets an individual TDM circuit. Another alternative is for the service provider to introduce packet multiplexing in the access switch. The service provider can achieve cost savings by having multiple customers share the same TDM circuit. These cost savings translate into lower cost for the basic connectivity service provided to the customer.

Figure 2-9 shows a scenario where, in the same 50-building metro, each building has an STS-1 (DS3) link that is shared by all 20 customers in each building. This greatly reduces the number of TDM circuits that have to be provisioned, because all customers in the same building share the same STS-1 circuit toward the CO. This reduces the total TDM circuits from 1000 to 50. Notice that 50 Ethernet ports still need to be terminated in the CO if the cross-connect does not have tagging or packet-multiplexing capabilities.

**Figure 2-9** EOS with Packet Multiplexing at the Access

Packet multiplexing in the last mile is yet another incremental step that the metro carriers would have to adopt to move in the direction of delivering switched Ethernet services. Although this model reduces the number of TDM circuits that need to be provisioned, it introduces issues of circuit oversubscription and SLA guarantees. Traffic from multiple customers would be fighting for the STS-1 link, and one customer with a Fast Ethernet (100 Mbps) interface could easily oversubscribe the 45-Mbps link. The carrier would need to use techniques such as traffic policing and traffic shaping to be able to sell its customers tiered bandwidth. These techniques are discussed in Chapter 3, “Metro Ethernet Services,” as part of a discussion about bandwidth parameters defined by the Metro Ethernet Forum (MEF).

## EOS with Packet Switching

The discussion thus far has addressed the ability to deliver a basic point-to-point leased-line or Internet-connectivity service. More-advanced VPN services can also be delivered over a SONET/SDH metro network that supports EOS. With a VPN service, the assumption is that different locations of the “same” customer exist in a metro area, and the customer wants to be able to tie to these locations via a virtual network. This type of service requires packet switching. Of course, if all the customer wants is a point-to-point service, no switching is required.

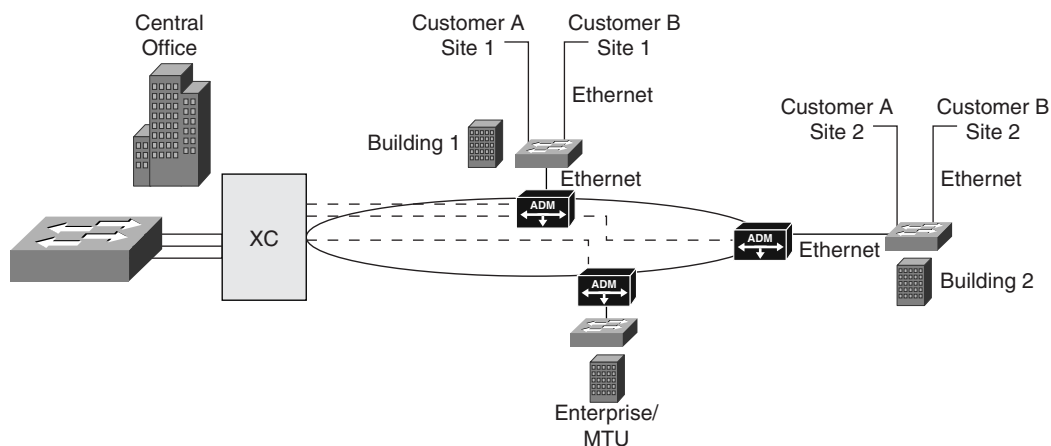
Packet switching can be delivered using either of two methods:

- Centralized switching
- Local switching

## EOS with Centralized Switching

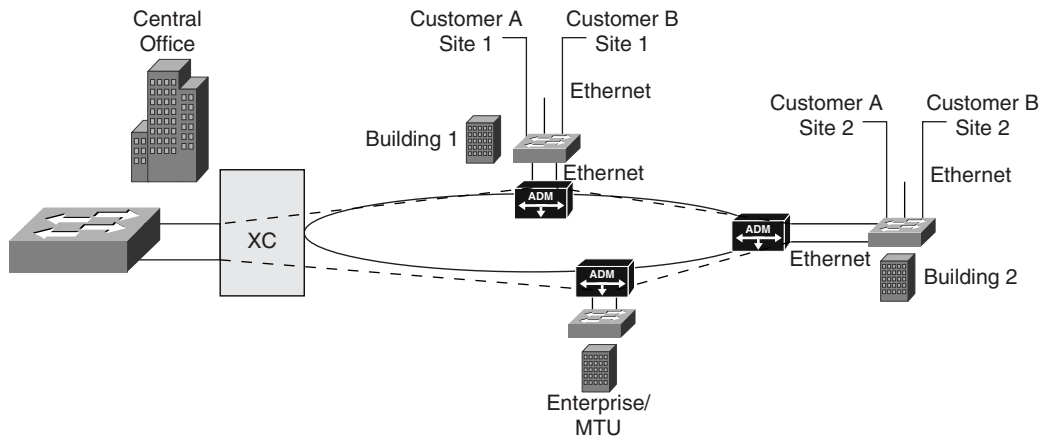
With centralized switching, a TDM circuit is provisioned from each building to the CO. All circuits are terminated in the CO, which is where the packet switching happens. Note that the standard SONET/SDH operation in unidirectional path switched rings (UPSRs) is to have active circuits and protect circuits on the other side of the ring to achieve the 50-ms ring failure. So, in the metro that has 50 buildings, 50 active STS-1s and 50 protect STS-1 circuits are provisioned. Also note that in case the XC does not support packet tagging or switching, 50 Ethernet interfaces need to be connected to the Ethernet switch at the CO. In Figure 2-10, customer A in sites 1 and 2 belongs to VPN A, while customer B in sites 1 and 2 belongs to VPN B.

**Figure 2-10** *EOS with Centralized Switching*

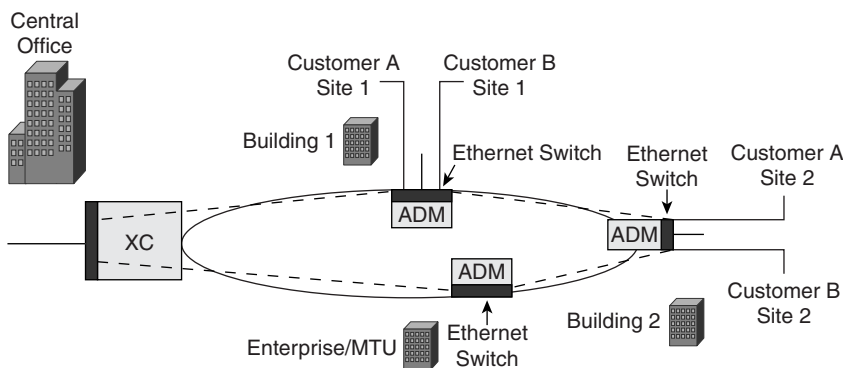


## EOS with Local Switching

With local switching, packet switching occurs on each node in the ring. The difference here is that TDM circuits are no longer provisioned between the buildings and the CO but are instead provisioned around the ring. Each ADM in the building terminates circuits for both east and west, and packets get switched at the local switching function in the basement of the building, as shown in Figure 2-11. In this case, SONET/SDH ring protection is not used. The metro carrier must rely on higher-level protection. In the case of L2 Ethernet, this means implementing standard spanning-tree mechanisms, such as the Spanning Tree Protocol (STP), or some other proprietary mechanisms that the Ethernet switch vendor offers. For example, STP would block one side of the ring to prevent a broadcast storm. Also note in Figure 2-11 that in each ADM, a separate Ethernet interface is dedicated to each TDM circuit that gets terminated, unless the ADM itself has a packet-switching function to aggregate the traffic toward the building. If more bandwidth is needed for the building, VCAT can be used to aggregate more circuits while still making them look like a single pipe.

**Figure 2-11** EOS with Local Switching

A variation of local switching is to integrate the Ethernet switching function and the ADM/EOS function into one box, as shown in Figure 2-12.

**Figure 2-12** A Variation of EOS with Local Switching

In this case, the TDM circuits are still terminated at each switch/ADM box on the ring. The benefit of this model is that it reduces the number of boxes deployed in the network; however, it blurs the line between the operation of data and TDM networks.

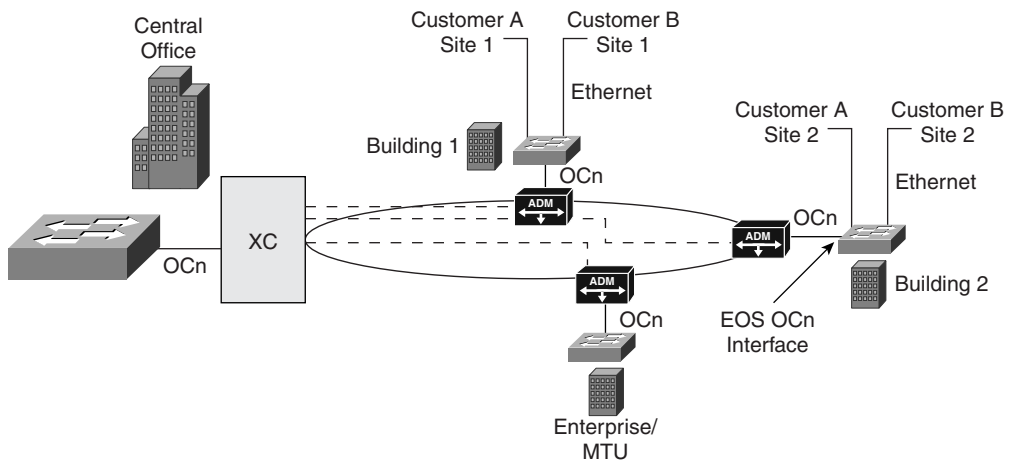
You probably realize by now why metro carriers that are used to SONET/SDH provisioning would like to stay close to the same old circuit-provisioning model. The terms “spanning tree” and “broadcast storms” give metro operators the jitters, because these are enterprise terms that sound very threatening for carriers that are bound to strict SLAs.

## EOS Interfaces in the Data Equipment

So far, this chapter has discussed different scenarios for having an EOS interface within the transport equipment. This section discusses the scenario in which the EOS interfaces are part of the data equipment rather than the transport equipment. In this model, the transport equipment does not have to deal with mapping the Ethernet frames carried in the SONET/SDH payload; the data-switching equipment does that instead.

The EOS interfaces inside the data equipment, as shown in Figure 2-13, are SONET/SDH interfaces with a mapping function that maps the EOS frames carried inside the SONET/SDH payload to an Ethernet frame. The Ethernet frame is in turn presented to the switching logic inside the data equipment. As in the case of transport equipment, the EOS interface can support VCAT. The advantage of this model is that the switching function, the EOS function, and the VCAT functions are all in the same box and are decoupled from the TDM box, which may already be installed in the network. This allows the data equipment to have better control over mapping the different data streams over the SONET/SDH circuits. With this model, multipoint-to-multipoint switched Ethernet services can be delivered efficiently while leveraging the existing legacy SONET/SDH infrastructure. This also fits better with the current operational model, in which transport and data are managed separately.

**Figure 2-13** *EOS in the Data Equipment*



The previously mentioned EOS scenarios are bound to create a lot of debates and confusion in the industry. From a technology perspective, all options are viable, assuming the vendor equipment is capable of delivering the services. From a business perspective, the ownership of the EOS interface determines who makes money on the sale: the data-switching vendors or the transport vendors. You will see numerous debates from both ends about where the EOS services and functions such as VCAT start and terminate.



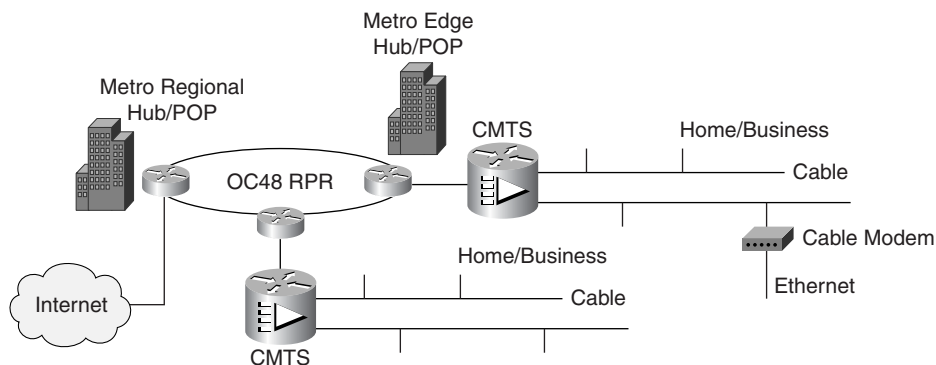
## Resilient Packet Ring

RPR also plays an important role in the development of data services in the metro. RPR is a new Media Access Control (MAC) protocol that is designed to optimize bandwidth management and to facilitate the deployment of data services over a ring network. The roots of RPR go back to the point at which Cisco Systems adopted a proprietary Data Packet Transport (DPT) technology to optimize packet rings for resiliency and bandwidth management. DPT found its way into the IEEE 802.17 workgroup, which led to the creation of an RPR standard that differs from the initial DPT approach.

RPR has so far been a very attractive approach to multiple service operators (MSOs), such as cable operators that are aggregating traffic from cable modem termination systems (CMTSs) in the metro. It remains to be seen whether RPR will be deployed by the incumbent carriers, such as the RBOCs and ILECs, that so far haven't been widely attracted to the RPR concept. The primary reason why they lack interest is that they view RPR deployments as new deployments, compared to EOS deployments, which leverage existing infrastructure and are therefore more evolutionary. RPR is a new packet-ring technology that is deployed over dark fiber or wavelength division multiplexing (WDM) instead of the traditional SONET/SDH rings. RPR could be deployed as an overlay over existing SONET/SDH infrastructure; however, the complexity of overlaying logical RPR rings over physical SONET/SDH rings will probably not be too attractive to many operators. Although RPR and EOS solve different issues in the metro (EOS solves Ethernet service deployment, and RPR solves bandwidth efficiency on packet rings), both technologies will compete for the metro provider's mind share.

Figure 2-14 shows a typical RPR deployment with a cable operator. The CMTSs aggregate the traffic coming over the coaxial cable from businesses and homes and hand over the data portion (assuming the cable is carrying voice/video as well) to the RPR router. Multiple RPR routers connect via an OC48 packet ring, and the traffic gets aggregated in the core hub, where Internet connectivity is established.

**Figure 2-14** *RPR Deployments*



RPR is somehow more commonly associated with routers than with switches, whereas EOS is more commonly associated with switches than routers. The reason for these associations is that DPT historically has been deployed using Cisco IP routers for delivering routed IP services over a packet ring. While the IEEE 802.17 standards body would like to make RPR independent of Layer 2 (L2) switching or Layer 3 (L3) routing, the fact remains that RPR has so far been adopted for L3 services. Also, many routers lack the right functionality to deliver L2 services, which makes EOS more suitable for switches. Again, while the technologically savvy reader might argue that L2 or L3 could be delivered with either technology—and there are existing platforms that support both L2 and L3 services—service provider adoption will be the determining factor in how each technology will most likely be used.

In comparing RPR with traditional SONET/SDH rings, you will realize that RPR deployments have many advantages simply because RPR is a protocol built from the ground up to support data rings. The following sections discuss several features of RPR.

## RPR Packet Add, Drop, and Forward

The RPR operation consists of three basic operations: add, drop, and forward. These operations mimic the add/drop mechanisms that are used in traditional SONET networks, where circuits get added, dropped, and cross-connected inside a ring.

The advantage that RPR has over a traditional Ethernet switched packet ring is that Ethernet 802.3 MAC operation processes packets at each node of the ring irrespective of whether the packet destination is behind that node. In contrast, RPR 802.17 MAC forwards the traffic on the ring without doing any intermediary switching or buffering if the traffic does not belong to the node. This reduces the amount of work individual nodes have to do.

In the RPR operation shown in Figure 2-15, traffic that does not belong to a particular node is transited (forwarded) on the ring by the 802.17 MAC. In the Ethernet 802.3 MAC operation, the traffic is processed and buffered at each node for the switching function to determine the exit interface.

RPR's advantage over a SONET/SDH ring is that all the packets coming into the ring share the full-ring bandwidth, and the RPR mechanism manages the bandwidth allocation to avoid congestion and hot spots. In a SONET/SDH ring, TDM time slots are allocated to each circuit, and the bandwidth is removed from the ring whether there is traffic on that circuit or not.

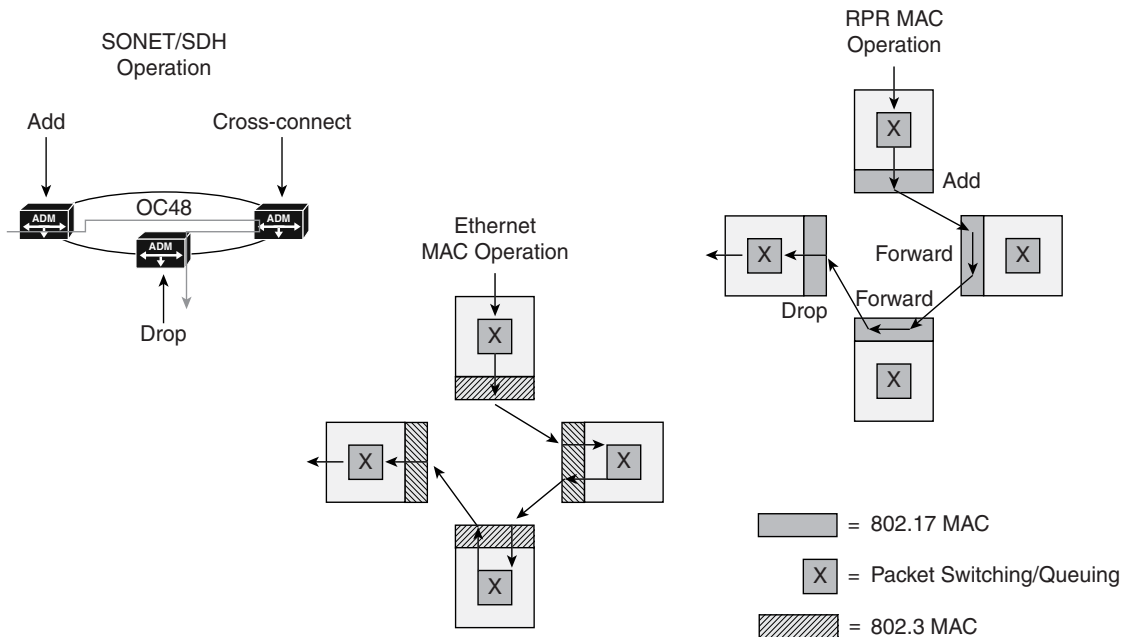
## RPR Resiliency

RPR offers ring protection in 50 ms, comparable with the traditional SONET/SDH protection. RPR fast protection with full-ring bandwidth utilization is probably one of the main assets that RPR has when compared to SONET/SDH and other packet-protection mechanisms.

RPR protection is achieved in two ways:

- **Ring wrapping**—The ring is patched at the location of the fault.
- **Ring steering**—In case of failure, the traffic is redirected (steered) at the source toward the working portion of the ring.

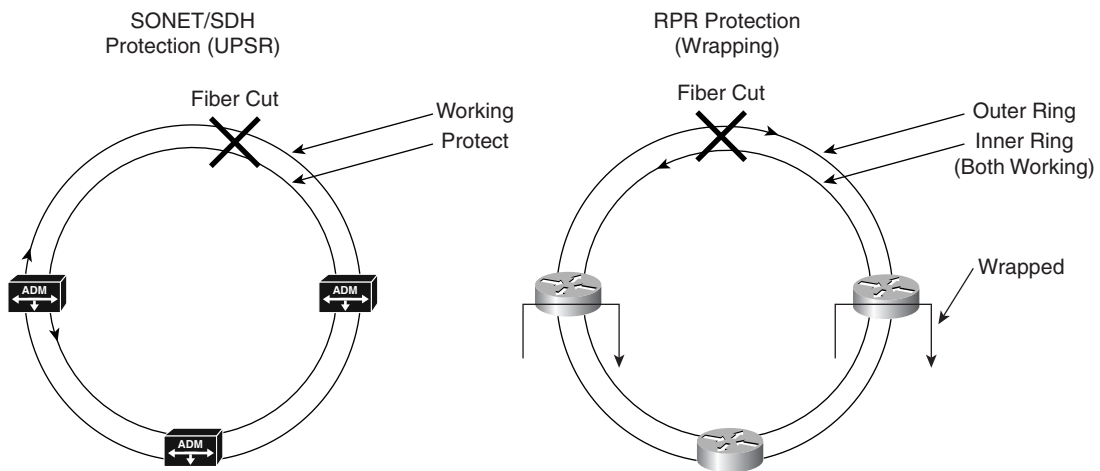
**Figure 2-15** RPR Add, Drop, and Forward



In general, the physical layer detects faults and signals that information to the MAC layer. If the failure is determined to be critical, each affected RPR node initiates a fail-over action for the service flows it originates that are affected by the facility outage. The fail-over action is a simple redirection of the traffic from the failed path to the protection path. The process of alarm notification and redirecting traffic is completed within 50 ms of the outage.

Figure 2-16 compares and contrasts RPR and SONET/SDH. In the SONET/SDH UPSR schemes, for example, 50-ms protection is achieved by having a working fiber and a standby protect fiber at all times. A sending node transmits on both fibers (east and west) at the same time, and a receiving node accepts traffic from only one side. In case of a fiber cut, recovery is done in less than 50 ms. In UPSRs, only 50 percent of the fiber capacity is used, because the other half is kept for failure mode. In RPR, both fiber rings—the outer ring and the inner ring—are used to utilize 100 percent of the ring capacity. In case of a failure, the ring wraps, isolating the failed portion. So, in essence, the effective bandwidth of an RPR ring is twice as much as a SONET/SDH ring because of the SONET/SDH protection.

Figure 2-16 RPR Protection



## RPR Fairness

RPR implements a fairness algorithm to give every node on the ring a fair share of the ring. RPR uses access-control mechanisms to ensure fairness and to bound latency on the ring. The access control can be broken into two types, which can be applied at the same time:

- **Global access control**—Controls access so that every node can get a fair share of the ring's global bandwidth.
- **Local access control**—Gives the node additional ring access—that is, bandwidth beyond what was globally allocated—to take advantage of segments that are less-used.

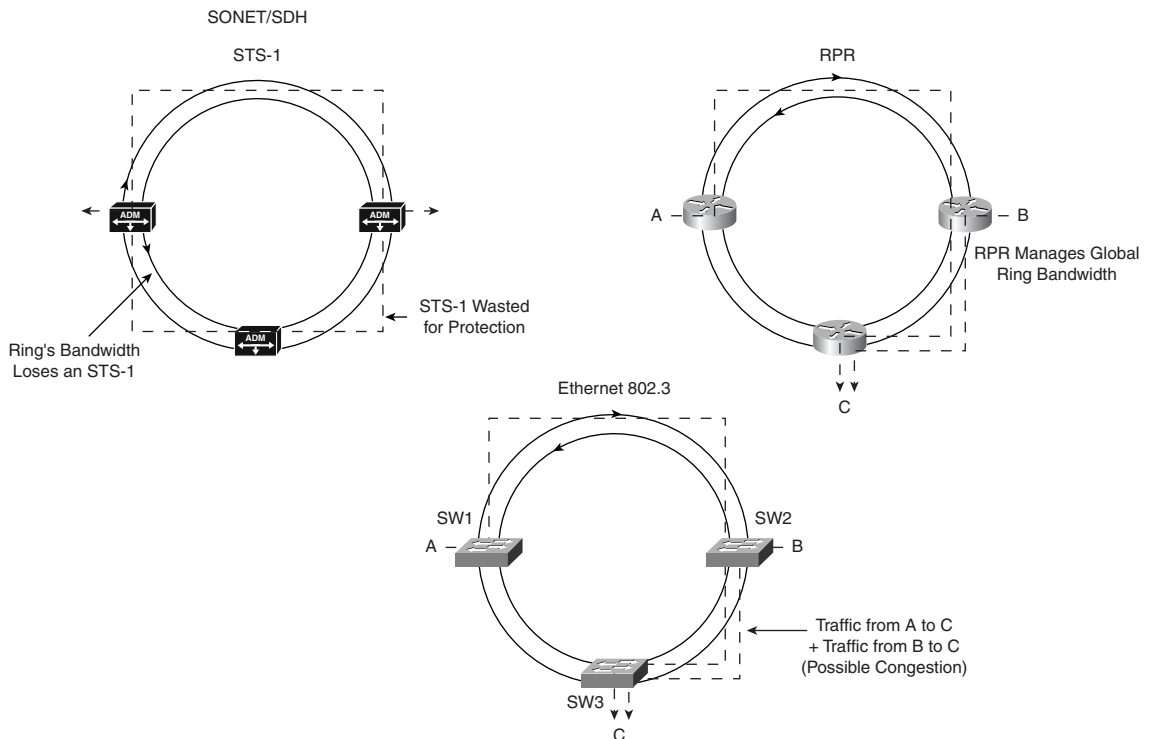
RPR uses the special reuse protocol (SRP), which is a concept used in rings to increase the ring's overall aggregate bandwidth. This is possible because multiple spans of the ring can be used simultaneously without having the traffic on one span affect the traffic on the other spans. If a node experiences congestion, it notifies the upstream nodes on the ring, which in turn adjust the transmit rate to relieve downstream congestion.

It helps to contrast ring bandwidth fairness between RPR and L2 Ethernet rings. In the case of an Ethernet ring with L2 switching, there is no such thing as ring fairness, because the QoS decisions are local to each node, irrespective of what is on the ring. You can use rate-limiting techniques to prevent a set of customers who are coming in on one node from oversubscribing the ring; however, it would be hard to have a global fairness mechanism without resorting to complicated QoS management software applications that would coordinate between all nodes.

Figure 2-17 shows three different scenarios for SONET/SDH UPSR, RPR, and L2 Ethernet rings. In the SONET/SDH case, if an STS-1 is allocated, the ring loses an STS-1 worth of

bandwidth, irrespective of actual traffic. In the Ethernet case, traffic from A to C and from B to C might oversubscribe the capacity of the point-to-point link between switches SW2 and SW3. In the RPR case, the MAC entity on each node monitors the utilization on its immediate links and makes that information available to all nodes on the ring. Each node can then send more data or throttle back.

**Figure 2-17** *Ring Bandwidth*



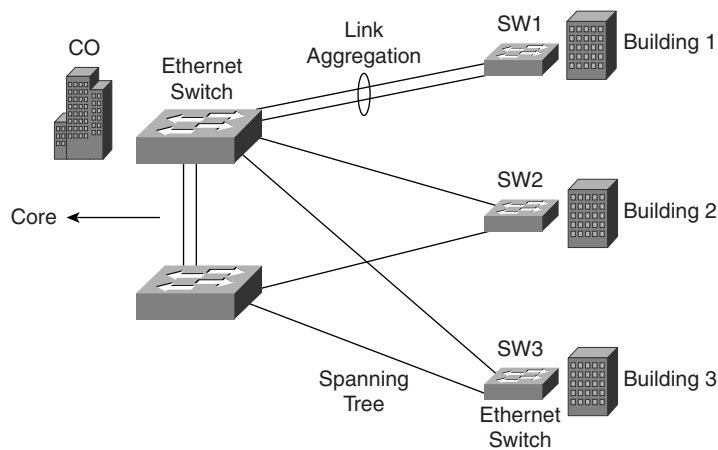
## Ethernet Transport

So far, this book has addressed the reasoning behind adopting Ethernet as an access interface rather than a TDM interface. But as discussed in this section, Ethernet isn't limited to being an access technology. Many efforts have been made to extend Ethernet itself into the MAN as a transport technology. Since the early 2000s, metro Ethernet deployments have taken many shapes and forms; some have proven to work, and others have not. When Ethernet is used as a transport technology, the access network can be built in either ring or hub-and-spoke topologies. These are discussed next.

## Gigabit Ethernet Hub-and-Spoke Configuration

In a Gigabit Ethernet hub-and-spoke configuration, Ethernet switches deployed in the basement of buildings are dual-homed into the nearest point of presence (POP) or CO. Dedicated fiber, or dedicated wavelengths using WDM, is used for connectivity. Although this is the most expensive approach for metro access deployments because of the cost of fiber, some carriers consider it to be the better solution as far as survivability and scalability compared to deploying Ethernet in a ring topology (described in the next section). With the hub-and-spoke model, the bandwidth dedicated to each building can scale, because the full fiber is dedicated to the building. Protection schemes can be achieved via mechanisms such as link aggregation 802.3ad or dual homing. With link aggregation, two fibers are aggregated into a bigger pipe that connects to the CO. Traffic is load-balanced between the two fibers, and if one fiber is damaged, the other absorbs the full load. This, of course, assumes that the two fibers are run into two different conduits to the CO for better protection. This scenario is shown in Figure 2-18 for the connection between building 1 and the CO.

**Figure 2-18** *Ethernet Hub and Spoke*



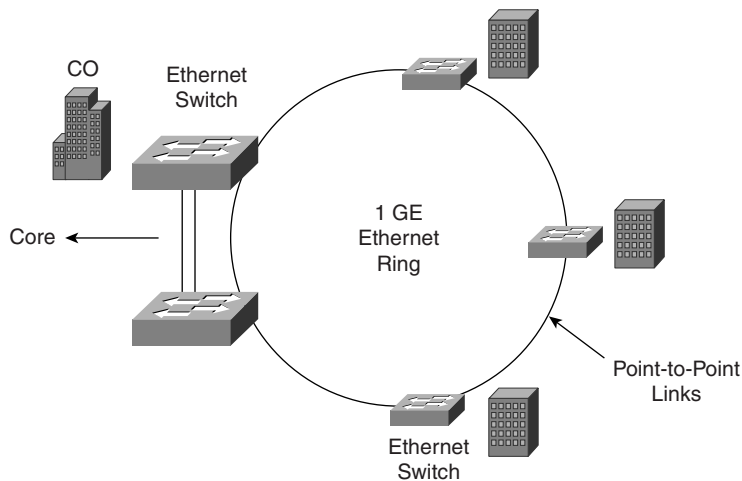
Another approach is to dual-home the fiber into different switches at the CO, as shown in Figure 2-18 for buildings 2 and 3. Although this prevents a single point of failure on the switching side, it creates more complexities, because STP must be run between the buildings and the CO, causing traffic on one of the dual-homed links to be blocked.

## Gigabit Ethernet Rings

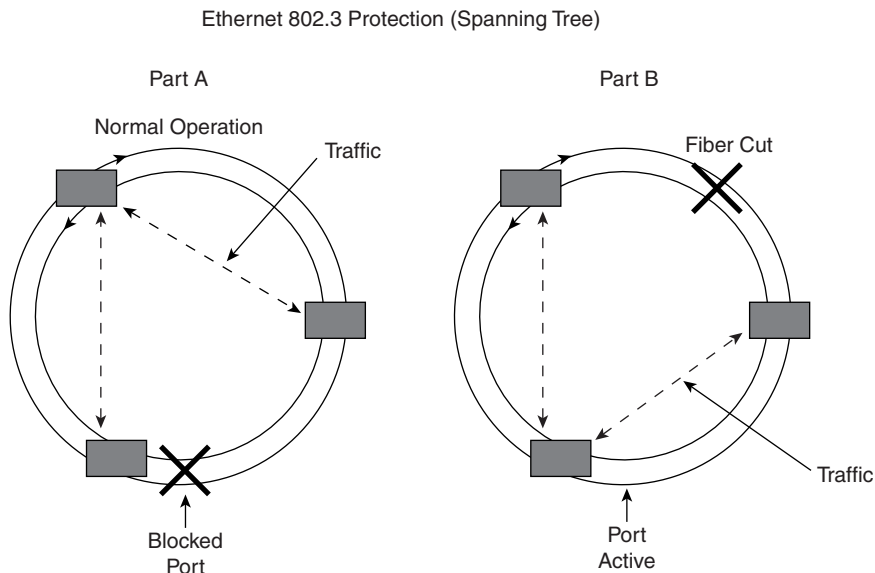
Many fiber deployments in the metro are laid in ring configurations. Consequently, ring topologies are the most natural to implement and result in cost savings. However, the situation differs depending on whether you are dealing with U.S. carriers or international carriers, incumbents, or greenfields. Ring deployments could be extremely cost-effective for one carrier

but a nonissue for another. For existing fiber laid out in a ring topology, Gigabit Ethernet rings are a series of point-to-point connections between the switches in the building basements and the CO, as shown in Figure 2-19. As simple as they might look, Gigabit Ethernet rings may create many issues for the operators because of protection and bandwidth limitations. First of all, ring capacity could be a major issue. Gigabit Ethernet rings have only 1 GB of capacity to share between all buildings, and some of that capacity is not available because spanning tree blocks portions of the ring to prevent loops.

**Figure 2-19** *Gigabit Ethernet Rings*



With an Ethernet L2 switched operation, the ring itself becomes a collection of point-to-point links. Even without a fiber cut, spanning tree blocks portions of the ring to prevent broadcast storms caused by loops (see Part A of Figure 2-20). Broadcast storms occur, for example, when a packet with an unknown destination reaches a node. The node floods the packet over the ring according to standard bridging operation as defined in 802.3d. If there is a loop in the network (in this case, the ring), the packet could end up being received and forwarded by the same node over and over. The spanning-tree algorithm uses control packets called bridge protocol data units (BPDUs) to discover loops and block them. Spanning tree normally takes between 30 and 60 seconds to converge. The new 802.1W Rapid Spanning Tree allows faster convergence but still doesn't come close to 50 ms. Many proprietary algorithms have been introduced to achieve ring convergence in less than 1 second, which many operators view as good enough for data services and even for Voice over IP (VoIP) services. However, because L2 switching cannot operate in a loop environment, many of these algorithms still need to block redundant paths in the ring to prevent broadcast storms, and are not considered as reliable as RPR or SONET/SDH protection mechanisms that are more carrier-class. When a fiber cut occurs, spanning tree readjusts, and the new path between the different nodes is established, as shown in Part B of Figure 2-20.

**Figure 2-20** Gigabit Ethernet Rings—Spanning Tree

Although 10-Gigabit Ethernet rings would alleviate the congestion issues, initial solutions for 10-GE switches are cost-prohibitive. Initial equipment with 10-GE interfaces was designed for core networks rather than building access. As 10-GE solutions mature and their prices are reduced to fit the building access, 10-GE rings will become a viable solution.



Other methods, such as deploying WDM, could be used to add capacity on the ring. It is debatable whether such methods are cost-effective for prime-time deployments, because they increase the operational overhead of deploying the access ring.

## Conclusion

So far, you have read about different technologies that can be used for physical metro connectivity. Ethernet over SONET, RPRs, and Ethernet transport are all viable methods to deploy a metro Ethernet service. Ethernet over SONET presents a viable solution for deploying Ethernet services over an existing installed base. You have seen how virtual concatenation allows better efficiency and bandwidth granularity when mapping Ethernet pipes over SONET/SDH rings. RPR is a packet-ring technology that is attracting much interest from MSOs because it solves many of the restoration and bandwidth inefficiencies that exist in SONET/SDH rings. Ethernet as a transport technology is also a simple and efficient way to deploy Ethernet services; however, by itself, this solution inherits many of the deficiencies of L2 switched Ethernet networks.



Much functionality still needs to be offered on top of metro equipment to deliver revenue-generating services such as Internet connectivity and VPN services. Ethernet, for example, has always been used in a single-customer environment, such as an enterprise network. It is now moving to a multicustomer environment in which the same equipment delivers services to different customers over a shared carrier network. Issues of virtualization of the service and service scalability become major issues. Ethernet over MPLS (EoMPLS) is becoming a viable solution for deploying scalable metro Ethernet services. The MPLS control plane delivers most of the functionality that is lacking in Ethernet switched networks as far as scalability and resiliency. Chapter 3 discusses metro Ethernet services and Layer 2 switching, in preparation for Chapter 4, which discusses delivering Ethernet over hybrid Ethernet and IP/MPLS networks.



This chapter covers the following topics:

- L2 Switching Basics
- Metro Ethernet Services Concepts
- Example of an L2 Metro Ethernet Service
- Challenges with All-Ethernet Metro Networks

# Metro Ethernet Services

---

As discussed in Chapter 1, “Introduction to Data in the Metro,” Ethernet services can take either of two forms: a retail service that competes with traditional T1/E1 private-line services, or a wholesale service where a carrier sells a big Ethernet transport pipe to another, smaller service provider. In either case, multiple customers share the same metro infrastructure and equipment. For TDM deployments, sharing the infrastructure is a nonissue, because the services are limited to selling transport pipes, and each customer is allocated a circuit that isolates its traffic from other customers. The customer gets well-defined SLAs, mainly dictated by the circuit that is purchased.

When packet multiplexing and switching are applied, such as in the cases of switched EOS, Ethernet Transport, and RPR, things change. Packets from different customers are multiplexed over the same pipe, and the bandwidth is shared. No physical boundaries separate one customer’s traffic from another’s, only logical boundaries. Separation of customer traffic and packet queuing techniques have to be used to ensure QoS. Multiple functions have to be well-defined to offer a service:

- How to identify different customers’ traffic over a shared pipe or shared network
- How to identify and enforce the service given to a particular customer
- How to allocate certain bandwidth to a specific customer
- How to “transparently” move customers’ traffic between different locations, such as in the case of transparent LAN services
- How to scale the number of customers
- How to deploy a VPN service that offers any-to-any connectivity for the same customer

This chapter starts by discussing the basics of L2 Ethernet switching to familiarize you with Ethernet-switching concepts. Then it discusses the different metro Ethernet service concepts as introduced by the Metro Ethernet Forum (MEF).

## L2 Switching Basics

L2 switching allows packets to be switched in the network based on their Media Access Control (MAC) address. When a packet arrives at the switch, the switch checks the packet’s destination MAC address and, if known, sends the packet to the output port from which it learned the destination MAC.

The two fundamental elements in Ethernet L2 switching are the MAC address and the virtual LAN (VLAN). In the same way that IP routing references stations on the networks via an L3 IP address, Ethernet L2 switching references end stations via the MAC address. However, unlike IP, in which IP addresses are assigned by administrators and can be reused in different private networks, MAC addresses are supposed to be unique, because they are indicative of the hardware itself. Thus, MAC addresses should not be assigned by the network administrator. (Of course, in some cases the MAC addresses can be overwritten or duplicated, but this is not the norm.)

Ethernet is a broadcast medium. Without the concept of VLANs, a broadcast sent by a station on the LAN is sent to all physical segments of the switched LAN. The VLAN concept allows the segmentation of the LAN into logical entities, and traffic is localized within those logical entities. For example, a university campus can be allocated multiple VLANs—one dedicated for faculty, one dedicated for students, and the third dedicated for visitors. Broadcast traffic within each of these VLANs is isolated to that VLAN.

Figure 3-1 shows the concept of an Ethernet LAN using a hub (Part A) and an Ethernet switch (Part B). With an Ethernet hub, all stations on the LAN share the same physical segment. A 10-Mbps hub, for example, allows broadcast and unicast traffic between the stations that share the 10-Mbps bandwidth. The switched LAN on the right allows each segment a 100-Mbps connection (for this example), and it segments the LAN into two logical domains, VLAN 10 and VLAN 20. The concept of VLANs is independent of the stations themselves. The VLAN is an allocation by the switch. In this example, ports 1 and 2 are allocated to VLAN 10, while ports 3 and 4 are allocated to VLAN 20. When stations A1 and A2 send traffic, the switch tags the traffic with the VLAN assigned to the interface and makes the switching decisions based on that VLAN number. The result is that traffic within a VLAN is isolated from traffic within other VLANs.

Ethernet switching includes the following basic concepts:

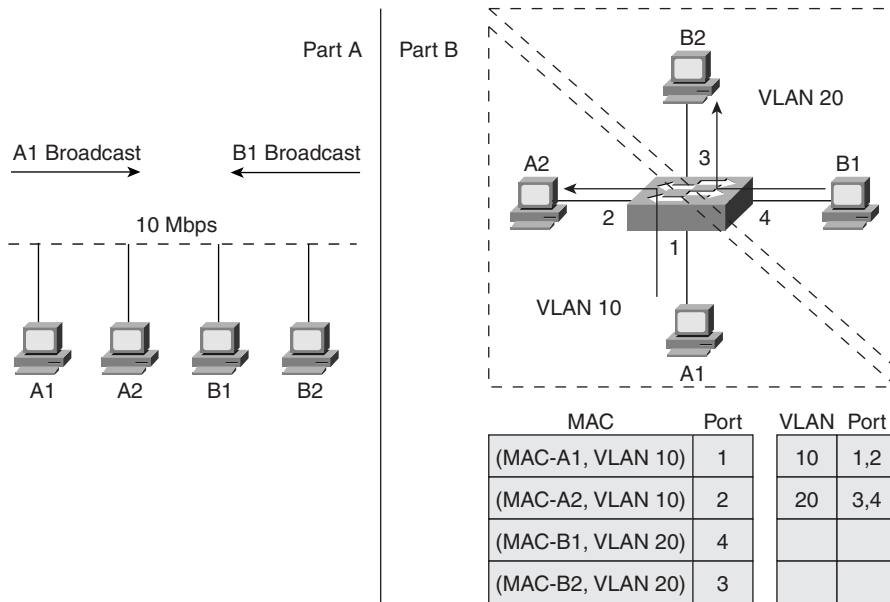
- MAC learning
- Flooding
- Using broadcast and multicast
- Expanding the network with trunks
- VLAN tagging
- The need for the Spanning Tree Protocol (STP)

## MAC Learning

MAC learning allows the Ethernet switch to learn the MAC addresses of the stations in the network to identify on which port to send the traffic. LAN switches normally keep a MAC learning table (or a bridge table) and a VLAN table. The MAC learning table associates the MACs/VLANs with a given port, and the VLAN table associates the port with a VLAN. In Figure 3-1, Part B, the switch has learned the MAC addresses of stations A1, A2, B1, and B2

on ports 1, 2, 4, and 3, respectively. It also shows that ports 1 and 2 are associated with VLAN 10 and ports 3 and 4 are associated with VLAN 20.

**Figure 3-1** Ethernet MACs and VLANs



## Flooding

If the switch receives a packet with a destination MAC address that does not exist in the bridge table, the switch sends that packet over all its interfaces that belong to the same VLAN assigned to the interface where the packet came in from. The switch does not flood the frame out the port that generated the original frame. This mechanism is called *flooding*. It allows the fast delivery of packets to their destinations even before all MAC addresses have been learned by all switches in the network. The drawback of flooding is that it consumes switch and network resources that otherwise wouldn't have been used if the switch had already learned which port to send the packet to.

VLANs minimize the effect of flooding because they concentrate the flooding within a particular VLAN. The switch uses the VLAN table to map the VLAN number of the port on which the packet arrived to a list of ports that the packet is flooded on.

## Using Broadcast and Multicast

Broadcast is used to enable clients to discover resources that are advertised by servers. When a server advertises its services to its clients, it sends broadcast messages to MAC address FFFF FFFF FFFF, which means "all stations." End clients pick up

JUNIPER Exhibit 1003

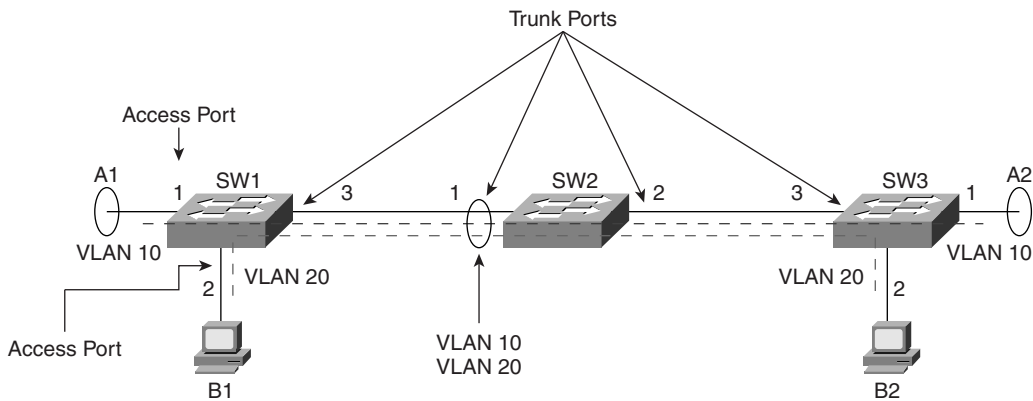
App. 6, pg. 66

only the broadcasts they are interested in, to minimize their CPU usage. With multicast, a subset of broadcast, a station sends traffic only to a group of stations and not to all stations. Broadcast and multicast addresses are treated as unknown destinations and are flooded over all ports within a VLAN. Some higher-layer protocols such as IGMP snooping help mitigate the flooding of IP multicast packets over an L2 switched network by identifying which set of ports a packet is to be flooded on.

## Expanding the Network with Trunks

So far you have seen the case of a single L2 switch. An L2 Ethernet-switched network would consist of many interconnected switches with trunk ports. The trunk ports are similar to the access ports used to connect end stations; however, they have the added task of carrying traffic coming in from many VLANs in the network. This scenario is shown in Figure 3-2. Trunk ports could connect Ethernet switches built by different vendors—hence the need for standardization for VLAN tagging mechanisms.

**Figure 3-2** *Trunk Ports*

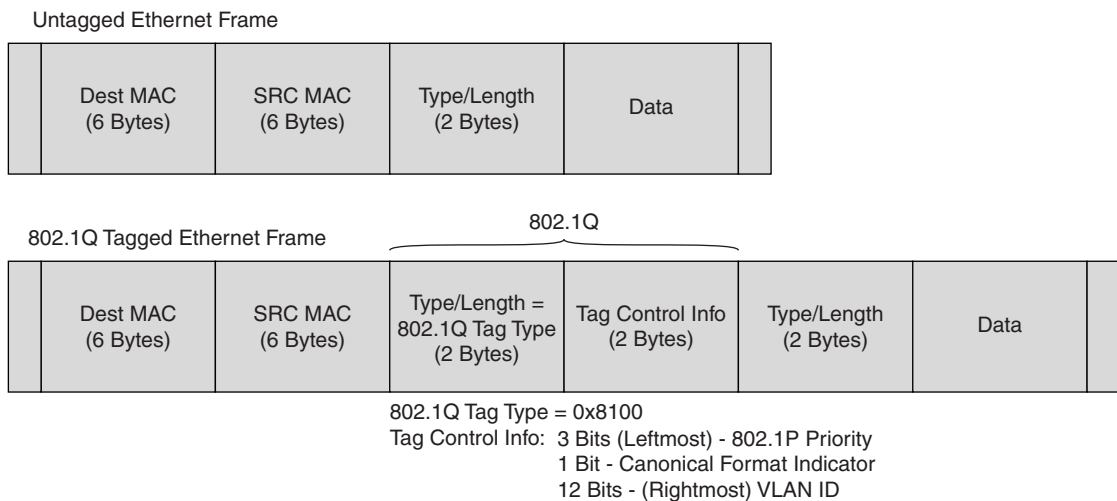


In Figure 3-2, switches SW1 and SW3 have assigned access port 1 with VLAN 10 and access port 2 with VLAN 20. Port 3 is a trunk port that is used to connect to other switches in the network. Note that SW2 in the middle has no access ports and is used only to interconnect trunk ports. You can see that the simplicity of switched Ethernet becomes extremely complex because VLAN assignments need to be tracked inside the network to allow the right traffic to be switched on the right ports. In Frame Relay, ATM, and MPLS, similar complexities do exist, and signaling is introduced to solve the network connectivity issues. Ethernet has *not* defined a signaling protocol. The only mechanisms that Ethernet networks have are third-party applications that surf the network and make it easier to do some VLAN allocations. While these mechanisms work in small enterprise environments, they immediately became showstoppers in larger enterprise deployments and carrier networks. Chapter 4 discusses LDP as a signaling mechanism for delivering Ethernet services. Chapter 7 discusses RSVP-TE and its use in relation to scaling the Ethernet services.

## VLAN Tagging

IEEE 802.1Q defines how an Ethernet frame gets tagged with a VLAN ID. The VLAN ID is assigned by the switch and not the end station. The switch assigns a VLAN number to a port, and every packet received on that port gets allocated that VLAN ID. The Ethernet switches switch packets between the same VLANs. Traffic between different VLANs is sent to a routing function within the switch itself (if the switch supports L3 forwarding) or an external router. Figure 3-3 shows how the VLAN tags get inserted inside the untagged VLAN packet.

**Figure 3-3** VLAN Tagged Packet



The untagged Ethernet packet consists of the destination MAC address and source MAC address, a Type field, and the data. The 802.1Q tag header gets inserted between the source MAC address and the Type field. It consists of a 2-byte Type field and a 2-byte Tag Control Info field. The 2-byte Type field is set to 0X8100 to indicate an 802.1Q tagged packet. The 2-byte Tag Control Info field consists of the 3 leftmost bits indicating the 802.1P priority and the 12 rightmost bits indicating the VLAN tag ID. The 802.1P field gives the Ethernet packet up to eight different priority levels that can be used to offer different levels of service within the network. The 12-bit VLAN ID field allows the assignment of up to 4096 ( $2^{12}$ ) VLAN numbers to distinguish the different VLAN tagged packets.

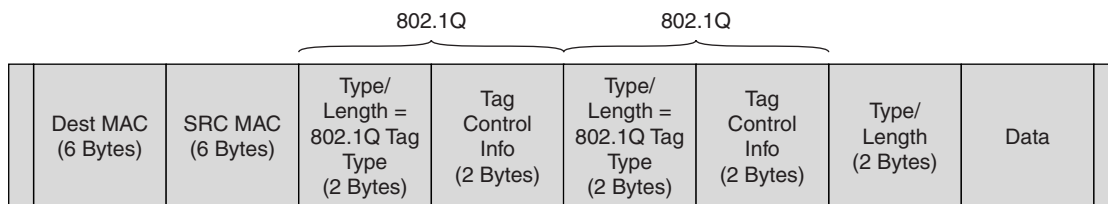
Metro Ethernet applications require extensions to L2 switching that are not defined in the standards. An example is the ability to do VLAN stacking—that is, to do multiple VLAN tagging to the same Ethernet packet and create a stack of VLAN IDs. Different entities can do L2 switching on the different levels of the VLAN stack. Cisco Systems calls this concept *Q-in-Q*, short for *802.1Q-in-802.1Q*, as shown in Figure 3-4.

As shown, an already tagged frame can be tagged again to create a hierarchy. The simplicity of Ethernet, the lack of standardization for many such extensions, the reliance on STP, and

the explosion of MAC addresses contribute to the lack of confidence of many providers in deploying a large-scale, all-Ethernet network.

**Figure 3-4** *Q-in-Q*

802.1Q-in-802.1Q (Q-in-Q) Tagged Ethernet Frame



VLAN tag support is discussed more in the section “VLAN Tag Support Attribute.”

## The Need for the Spanning Tree Protocol

L2 Ethernet-switched networks work on the basis of MAC address learning and flooding. If multiple paths exist to the same destination, and the packet has an unknown destination, packet flooding might cause the packet to be sent back to the original switch that put it on the network, causing a broadcast storm. STP prevents loops in the network by blocking redundant paths and ensuring that only one active path exists between every two switches in the network. STP uses bridge protocol data units (BPDUs), which are control packets that travel in the network and identify which path, and hence ports, need to be blocked.

The next section covers in detail the Ethernet services concepts as defined by the Metro Ethernet Forum.

## Metro Ethernet Services Concepts

The Metro Ethernet Forum is a nonprofit organization that has been active in defining the scope, concepts, and terminology for deploying Ethernet services in the metro. Other standards bodies, such as the Internet Engineering Task Force (IETF), have also defined ways of scaling Ethernet services through the use of MPLS. While the terminologies might differ slightly, the concepts and directions taken by these different bodies are converging.

For Ethernet services, the MEF defines a set of attributes and parameters that describe the service and SLA that are set between the metro carrier and its customer.

## Ethernet Service Definition

The MEF defines a User-to-Network Interface (UNI) and Ethernet Virtual Connection (EVC). The UNI is a standard Ethernet interface that is the point of demarcation between the customer equipment and the service provider’s metro Ethernet network.



The EVC is defined by the MEF as “an association of two or more UNIs.” In other words, the EVC is a logical tunnel that connects two (P2P) or more (MP2MP) sites, enabling the transfer of Ethernet frames between them. The EVC also acts as a separation between the different customers and provides data privacy and security similar to Frame Relay or ATM permanent virtual circuits (PVCs).

The MEF has defined two Ethernet service types:

- **Ethernet Line Service (ELS)**—This is basically a point-to-point (P2P) Ethernet service.
- **Ethernet LAN Service (E-LAN)**—This is a multipoint-to-multipoint (MP2MP) Ethernet service.

The Ethernet Line Service provides a P2P EVC between two subscribers, similar to a Frame Relay or private leased-line service (see Figure 3-5).

**Figure 3-5** *Ethernet Service Concepts*

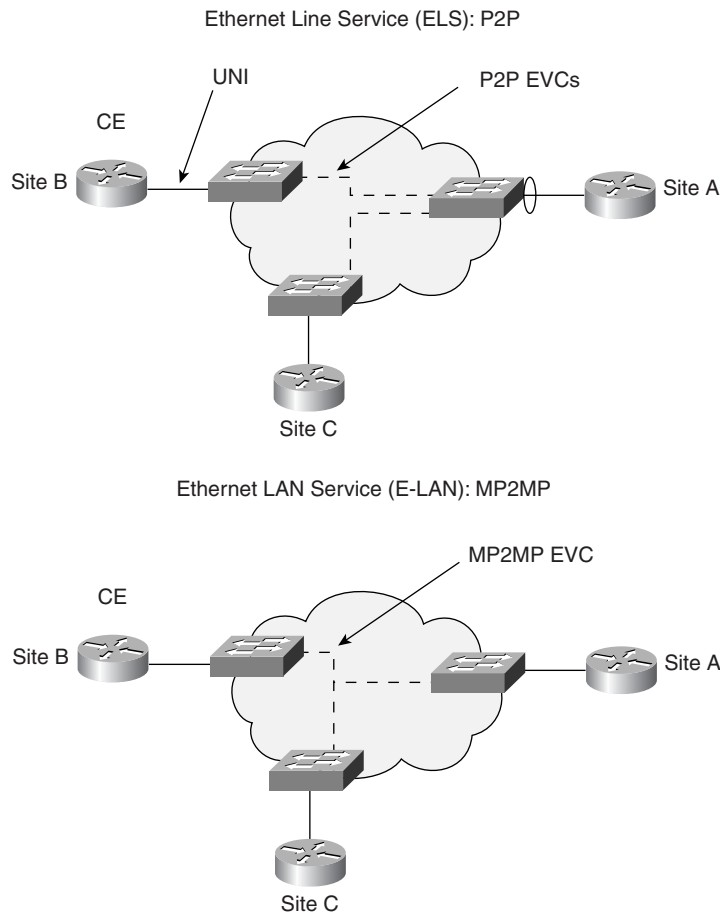


Figure 3-5 also illustrates the E-LAN, which provides multipoint connectivity between multiple subscribers in exactly the same manner as an Ethernet-switched network. An E-LAN service offers the most flexibility in providing a VPN service because one EVC touches all sites. If a new site is added to the VPN, the new site participates in the EVC and has automatic connectivity to all other sites.

## Ethernet Service Attributes and Parameters

The MEF has developed an Ethernet services framework to help subscribers and service providers have a common nomenclature when talking about the different service types and their attributes. For each of the two service types, ELS and E-LAN, the MEF has defined the following service attributes and their corresponding parameters that define the capabilities of the service type:

- Ethernet physical interface attribute
- Traffic parameters
- Performance parameters
- Class of service parameters
- Service frame delivery attribute
- VLAN tag support attribute
- Service multiplexing attribute
- Bundling attribute
- Security filters attribute

### Ethernet Physical Interface Attribute

The Ethernet physical interface attribute has the following parameters:

- **Physical medium**—Defines the physical medium per the IEEE 802.3 standard. Examples are 10BASE-T, 100BASE-T, and 1000BASE-X.
- **Speed**—Defines the Ethernet speed: 10 Mbps, 100 Mbps, 1 Gbps, or 10 Gbps.
- **Mode**—Indicates support for full duplex or half duplex and support for autospeed negotiation between Ethernet ports.
- **MAC layer**—Specifies which MAC layer is supported as specified in the 802.3-2002 standard.

### Traffic Parameters

The MEF has defined a set of bandwidth profiles that can be applied at the UNI or to an EVC. A bandwidth profile is a limit on the rate at which Ethernet frames can traverse the UNI or the

EVC. Administering the bandwidth profiles can be a tricky business. For P2P connections where there is a single EVC between two sites, it is easy to calculate a bandwidth profile coming in and out of the pipe. However, for the cases where a multipoint service is delivered and there is the possibility of having multiple EVCs on the same physical interface, it becomes difficult to determine the bandwidth profile of an EVC. In such cases, limiting the bandwidth profile per UNI might be more practical.

The Bandwidth Profile service attributes are as follows:

- Ingress and egress bandwidth profile per UNI
- Ingress and egress bandwidth profile per EVC
- Ingress and egress bandwidth profile per CoS identifier
- Ingress bandwidth profile per destination UNI per EVC
- Egress bandwidth profile per source UNI per EVC

The Bandwidth Profile service attributes consist of the following traffic parameters:

- **CIR (Committed Information Rate)**—This is the minimum guaranteed throughput that the network must deliver for the service under normal operating conditions. A service can support a CIR per VLAN on the UNI interface; however, the sum of all CIRs should not exceed the physical port speed. The CIR has an additional parameter associated with it called the Committed Burst Size (CBS). The CBS is the size up to which subscriber traffic is allowed to burst in profile and not be discarded or shaped. The in-profile frames are those that meet the CIR and CBS parameters. The CBS may be specified in KB or MB. If, for example, a subscriber is allocated a 3-Mbps CIR and a 500-KB CBS, the subscriber is guaranteed a minimum of 3 Mbps and can burst up to 500 KB of traffic and still remain within the SLA limits. If the traffic bursts above 500 KB, the traffic may be dropped or delayed.
- **PIR (Peak Information Rate)**—The PIR specifies the rate above the CIR at which traffic is allowed into the network and that may get delivered if the network is not congested. The PIR has an additional parameter associated with it called the Maximum Burst Size (MBS). The MBS is the size up to which the traffic is allowed to burst without being discarded. The MBS can be specified in KB or MB, similar to CBS. A sample service may provide a 3-Mbps CIR, 500-KB CBS, 10-Mbps PIR, and 1-MB MBS. In this case, the following behavior occurs:
  - Traffic is less than or equal to CIR (3 Mbps)—Traffic is in profile with a guaranteed delivery. Traffic is also in profile if it bursts to CBS (500 KB) and may be dropped or delayed if it bursts beyond 500 KB.
  - Traffic is more than CIR (3 Mbps) and less than PIR (10 Mbps)—Traffic is out of profile. It may get delivered if the network is not congested and the burst size is less than MBS (1 MB).
  - Traffic is more than PIR (10 Mbps)—Traffic is discarded.

## Performance Parameters

The performance parameters indicate the service quality experienced by the subscriber. They consist of the following:

- Availability
- Delay
- Jitter
- Loss

### Availability

Availability is specified by the following service attributes:

- **UNI Service Activation Time**—Specifies the time from when the new or modified service order is placed to the time service is activated and usable. Remember that the main value proposition that an Ethernet service claims is the ability to cut down the service activation time to hours versus months with respect to the traditional telco model.
- **UNI Mean Time to Restore (MTTR)**—Specifies the time it takes from when the UNI is unavailable to when it is restored. Unavailability can be caused by a failure such as a fiber cut.
- **EVC Service Activation Time**—Specifies the time from when a new or modified service order is placed to when the service is activated and usable. The EVC service activation time begins when all UNIs are activated. For a multipoint EVC, for example, the service is considered active when all UNIs are active and operational.
- **EVC Availability**—Specifies how often the subscriber's EVC meets or exceeds the delay, loss, and jitter service performance over the same measurement interval. If an EVC does not meet the performance criteria, it is considered unavailable.
- **EVC (MTTR)**—Specifies the time from when the EVC is unavailable to when it becomes available again. Many restoration mechanisms can be used on the physical layer (L1), the MAC layer (L2), or the network layer (L3).

### Delay

Delay is a critical parameter that significantly impacts the quality of service (QoS) for real-time applications. Delay has traditionally been specified in one direction as one-way delay or end-to-end delay. The delay between two sites in the metro is an accumulation of delays, starting from one UNI at one end, going through the metro network, and going through the UNI on the other end. The delay at the UNI is affected by the line rate at the UNI connection and the supported Ethernet frame size. For example, a UNI connection with 10 Mbps and 1518-byte frame size would cause 1.2 milliseconds (ms) of transmission delay ( $1518 * 8 / 10^6$ ).

The metro network itself introduces additional delays based on the network backbone speed and level of congestion. The delay performance is defined by the 95th percentile (95 percent)

of the delay of successfully delivered egress frames over a time interval. For example, a delay of 15 ms over 24 hours means that over a period of 24 hours, 95 percent of the “delivered” frames had a one-way delay of less than or equal to 15 ms.

The delay parameter is used in the following attributes:

- Ingress and egress bandwidth profile per CoS identifier (UNI service attribute)
- Class of service (EVC service attribute)

### Jitter

Jitter is another parameter that affects the service quality. Jitter is also known as delay variation. Jitter has a very adverse effect on real-time applications such as IP telephony. The jitter parameter is used in the following service attributes:

- Ingress and egress bandwidth profile per CoS identifier (UNI service attribute)
- Class of service (EVC service attribute)

### Loss

Loss indicates the percentage of Ethernet frames that are in-profile and that are not reliably delivered between UNIs over a time interval. On a P2P EVC, for example, if 100 frames have been sent from a UNI on one end and 90 frames that are in profile have been received on the other end, the loss would be  $(100 - 90) / 100 = 10\%$ . Loss can have adverse effects, depending on the application. Applications such as e-mail and HTTP web browser requests can tolerate more loss than VoIP, for example. The loss parameter is used in the following attributes:

- Ingress and egress bandwidth profile per CoS identifier (UNI service attribute)
- Class of service (EVC service attribute)

## Class of Service Parameters

Class of service (CoS) parameters can be defined for metro Ethernet subscribers based on various CoS identifiers, such as the following:

- **Physical port**—This is the simplest form of QoS that applies to the physical port of the UNI connection. All traffic that enters and exits the port receives the same CoS.
- **Source/destination MAC addresses**—This type of classification is used to give different types of service based on combinations of source and destination MAC addresses. While this model is very flexible, it is difficult to administer, depending on the service itself. If the customer premises equipment (CPE) at the ends of the connections are Layer 2 switches that are part of a LAN-to-LAN service, hundreds or thousands of MAC addresses might have to be monitored. On the other hand, if the CPEs are routers, the MAC addresses that are monitored are those of the router interfaces themselves. Hence, the MAC addresses are much more manageable.

- **VLAN ID**—This is a very practical way of assigning CoS if the subscriber has different services on the physical port where a service is defined by a VLAN ID (these would be the carrier-assigned VLANs).
- **802.1p value**—The 802.1p field allows the carrier to assign up to eight different levels of priorities to the customer traffic. Ethernet switches use this field to specify some basic forwarding priorities, such as that frames with priority number 7 get forwarded ahead of frames with priority number 6, and so on. This is one method that can be used to differentiate between VoIP traffic and regular traffic or between high-priority and best-effort traffic. In all practicality, service providers are unlikely to exceed two or three levels of priority, for the sake of manageability.
- **Diffserv/IP ToS**—The IP ToS field is a 3-bit field inside the IP packet that is used to provide eight different classes of service known as IP precedence. This field is similar to the 802.1p field if used for basic forwarding priorities; however, it is located inside the IP header rather than the Ethernet 802.1Q tag header. Diffserv has defined a more sophisticated CoS scheme than the simple forwarding priority scheme defined by ToS. Diffserv allows for 64 different CoS values, called Diffserv codepoints (DSCPs). Diffserv includes different per-hop behaviors (PHBs), such as Expedited Forwarding (EF) for a low delay, low-loss service, four classes of Assured Forwarding (AF) for bursty real-time and non-real-time services, Class Selector (CS) for some backward compatibility with IP ToS, and Default Forwarding (DF) for best-effort services.

Although Diffserv gives much more flexibility to configure CoS parameters, service providers are still constrained with the issue of manageability. This is similar to the airline QoS model. Although there are so many ways to arrange seats and who sits where and so many types of food service and luggage service to offer travelers, airlines can manage at most only three or four levels of service, such as economy, economy plus, business class, and first class. Beyond that, the overhead of maintaining these services and the SLAs associated with them becomes cost-prohibitive.

### Service Frame Delivery Attribute

Because the metro network behaves like a switched LAN, you must understand which frames need to flow over the network and which do not. On a typical LAN, the frames traversing the network could be data frames or control frames. Some Ethernet services support delivery of all types of Ethernet protocol data units (PDUs); others may not. To ensure the full functionality of the subscriber network, it is important to have an agreement between the subscriber and the metro carriers on which frames get carried. The EVC service attribute can define whether a particular frame is discarded, delivered unconditionally, or delivered conditionally for each ordered UNI pair. The different possibilities of the Ethernet data frames are as follows:

- **Unicast frames**—These are frames that have a specified destination MAC address. If the destination MAC address is known by the network, the frame gets delivered to the exact destination. If the MAC address is unknown, the LAN behavior is to flood the frame within the particular VLAN.

- **Multicast frames**—These are frames that are transmitted to a select group of destinations. This would be any frame with the least significant bit (LSB) of the destination address set to 1, except for broadcast, where all bits of the MAC destination address are set to 1.
- **Broadcast frames**—IEEE 802.3 defines the broadcast address as a destination MAC address, FF-FF-FF-FF-FF-FF.

Layer 2 Control Processing packets are the different L2 control-protocol packets needed for specific applications. For example, BPDU packets are needed for STP. The provider might decide to tunnel or discard these packets over the EVC, depending on the service. The following is a list of currently standardized L2 protocols that can flow over an EVC:

- **IEEE 802.3x MAC control frames**—802.3.x is an XON/XOFF flow-control mechanism that lets an Ethernet interface send a PAUSE frame in case of traffic congestion on the egress of the Ethernet switch. The 802.3x MAC control frames have destination address 01-80-C2-00-00-01.
- **Link Aggregation Control Protocol (LACP)**—This protocol allows the dynamic bundling of multiple Ethernet interfaces between two switches to form an aggregate bigger pipe. The destination MAC address for these control frames is 01-80-C2-00-00-02.
- **IEEE 802.1x port authentication**—This protocol allows a user (an Ethernet port) to be authenticated into the network via a back-end server, such as a RADIUS server. The destination MAC address is 01-80-C2-00-00-03.
- **Generic Attribute Registration Protocol (GARP)**—The destination MAC address is 01-80-C2-00-00-2X.
- **STP**—The destination MAC address is 01-80-C2-00-00-00.
- **All-bridge multicast**—The destination MAC address is 01-80-C2-00-00-10.

## VLAN Tag Support Attribute

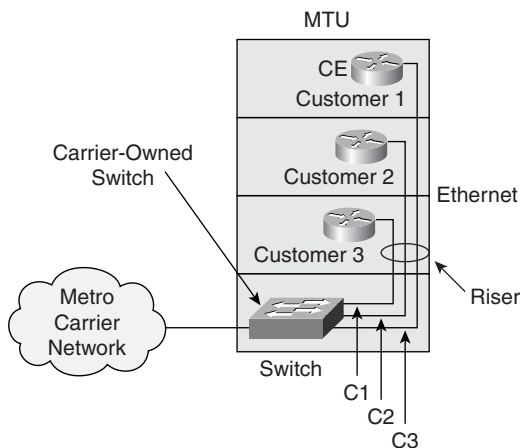
VLAN tag support provides another set of capabilities that are important for service frame delivery. Enterprise LANs are single-customer environments, meaning that the end users belong to a single organization. VLAN tags within an organization are indicative of different logical broadcast domains, such as different workgroups. Metro Ethernet creates a different environment in which the Ethernet network supports multiple enterprise networks that share the same infrastructure, and in which each enterprise network can still have its own segmentation. Support for different levels of VLANs and the ability to manipulate VLAN tags become very important.

Consider the example of an MTU building in which the metro provider installs a switch in the basement that offers multiple Ethernet connections to different small offices in the building. In this case, from a carrier perspective, each customer is identified by the physical Ethernet interface port that the customer connects to. This is shown in Figure 3-6.

Although identifying the customer itself is easy, isolating the traffic between the different customers becomes an interesting issue and requires some attention on the provider's part. Without special attention, traffic might get exchanged between customers in the

building through the basement switch. You have already seen in the section “L2 Switching Basics” that VLANs can be used to separate physical segments into many logical segments; however, this works in a single-customer environment, where the VLAN has a global meaning. In a multicustomer environment, each customer can have its own set of VLANs that overlap with VLANs from another customer. To work in this environment, carriers are adopting a model very similar to how Frame Relay and ATM services have been deployed. In essence, each customer is given service identifiers similar to Frame Relay data-link connection identifiers (DLCIs), which identify EVCs over which the customer’s traffic travels. In the case of Ethernet, the VLAN ID given by a carrier becomes that identifier. This is illustrated in Figure 3-7.

**Figure 3-6** *Ethernet in Multicustomer Environments*



In this example, the carrier needs to assign to each physical port a set of VLAN IDs that are representative of the services sold to each customer. Customer 1, for example, is assigned VLAN 10, customer 2 is assigned VLAN 20, and customer 3 is assigned VLAN 30. VLANs 10, 20, and 30 are carrier-assigned VLANs that are independent of the customer’s internal VLAN assignments. To make that distinction, the MEF has given the name CE-VLANs to the customer-internal VLANs. The customers themselves can have existing VLAN assignments (CE-VLANs) that overlap with each other and the carrier’s VLAN. There are two types of VLAN tag support:

- VLAN Tag Preservation/Stacking
- VLAN Tag Translation/Swapping

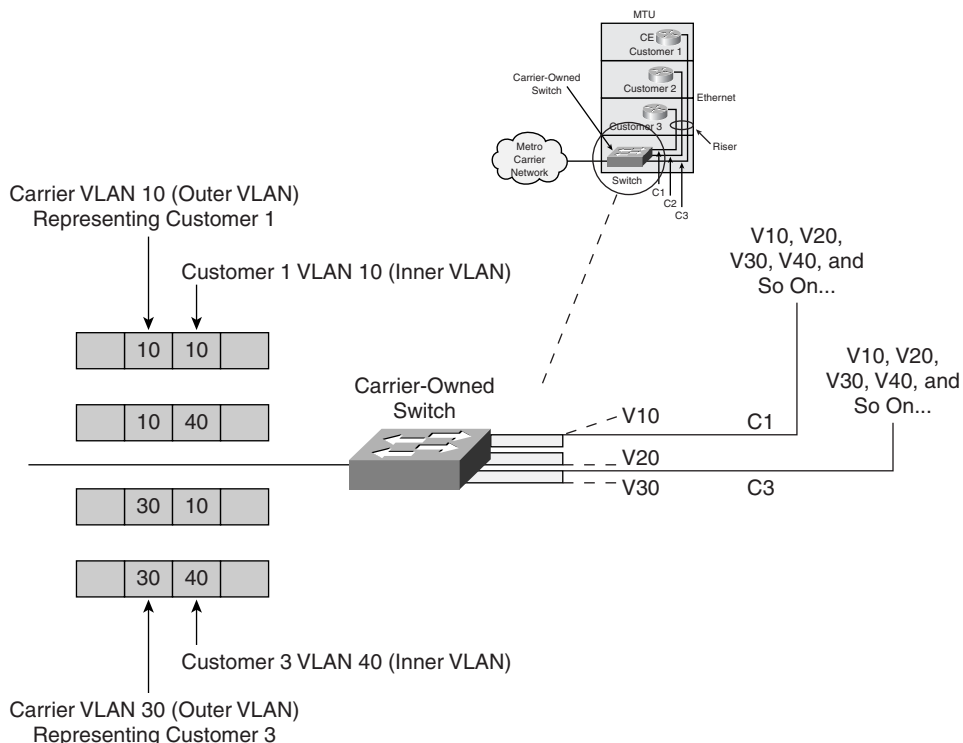
### VLAN Tag Preservation/Stacking

With VLAN Tag Preservation, all Ethernet frames received from the subscriber need to be carried untouched within the provider’s network across the EVC. This means that the VLAN ID at the ingress of the EVC is equal to the VLAN ID on the egress. This is typical of services such as LAN extension, where the same LAN is extended between two different locations and the enterprise-internal VLAN assignments need to be preserved. JUNIPER Exhibit 1003



switch supports multiple customers with overlapping CE-VLANs, the carrier's switch needs to be able to stack its own VLAN assignment on top of the customer's VLAN assignment to keep the separation between the traffic of different customers. This concept is called 802.1Q-in-802.1Q or Q-in-Q stacking, as explained earlier in the section "VLAN Tagging." With Q-in-Q, the carrier VLAN ID becomes indicative of the EVC, while the customer VLAN ID (CE-VLAN) is indicative of the internals of the customer network and is hidden from the carrier's network.

**Figure 3-7** Logical Separation of Traffic and Services



**WARNING** The Q-in-Q function is not standardized, and many vendors have their own variations. For the service to work, the Q-in-Q function must work on a “per-port” basis, meaning that each customer can be tagged with a different carrier VLAN tag. Some enterprise switches on the market can perform a double-tagging function; however, these switches can assign only a single VLAN-ID as a carrier ID for the whole switch. These types of switches work only if a single customer is serviced and the carrier wants to be able to carry the customer VLANs transparently within its network. These switches do not work when the carrier switch is servicing multiple customers, because it is impossible to differentiate between these customers using a single carrier VLAN tag.

### VLAN Tag Translation/Swapping

VLAN Tag Translation or Swapping occurs when the VLAN tags are local to the UNI, meaning that the VLAN tag value, if it exists on one side of the EVC, is independent of the VLAN tag values on the other side. In the case where one side of the EVC supports VLAN tagging and the other side doesn't, the carrier removes the VLAN tag from the Ethernet frames before they are delivered to the destination.

Another case is two organizations that have merged and want to tie their LANs together, but the internal VLAN assignments of each organization do not match. The provider can offer a service where the VLANs are removed from one side of the EVC and are translated to the correct VLANs on the other side of the EVC. Without this service, the only way to join the two organizations is via IP routing, which ignores the VLAN assignments and delivers the traffic based on IP addresses.

Another example of tag translation is a scenario where different customers are given Internet connectivity to an ISP. The carrier gives each customer a separate EVC. The carrier assigns its own VLAN-ID to the EVC and strips the VLAN tag before handing off the traffic to the ISP. This is illustrated in Figure 3-8.

**Figure 3-8** *VLAN Translation*

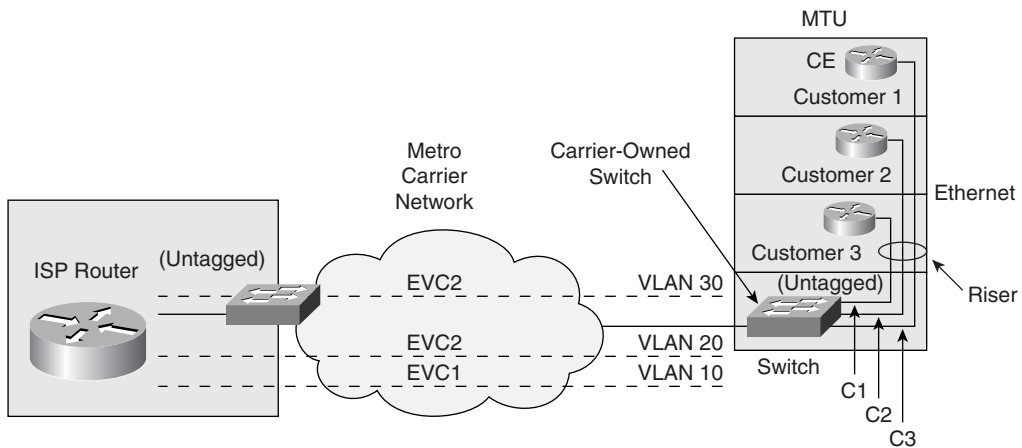


Figure 3-8 shows the metro carrier delivering Internet connectivity to three customers. The carrier is receiving untagged frames from the CPE routers located at each customer premises. The carrier inserts a VLAN tag 10 for all of customer 1's traffic, VLAN 20 for customer 2's traffic, and VLAN 30 for customer 3's traffic. The carrier uses the VLAN tags to separate the three customers' traffic within its own network. At the point of presence (POP), the VLAN tags are removed from all EVCs and handed off to an ISP router, which is offering the Internet IP service.

## Service Multiplexing Attribute

Service multiplexing is used to support multiple instances of EVCs on the same physical connection. This allows the same customer to have different services with the same Ethernet wire.

## Bundling Attribute

The Bundling service attribute enables two or more VLAN IDs to be mapped to a single EVC at a UNI. With bundling, the provider and subscriber must agree on the VLAN IDs used at the UNI and the mapping between each VLAN ID and a specific EVC. A special case of bundling is where every VLAN ID at the UNI interface maps to a single EVC. This service attribute is called *all-to-one bundling*.

## Security Filters Attribute

Security filters are MAC access lists that the carrier uses to block certain addresses from flowing over the EVC. This could be an additional service the carrier can offer at the request of the subscriber who would like a level of protection against certain MAC addresses. MAC addresses that match a certain access list could be dropped or allowed.

Tables 3-1 and 3-2 summarize the Ethernet service attributes and their associated parameters for UNI and EVCs.

**Table 3-1** *UNI Service Attributes*

UNI Service Attribute	Parameter Values or Range of Values
Physical medium	A standard Ethernet physical interface.
Speed	10 Mbps, 100 Mbps, 1 Gbps, or 10 Gbps.
Mode	Full-duplex or autospeed negotiation.
MAC layer	Ethernet and/or IEEE 802.3-2002.
Service multiplexing	Yes or no. If yes, all-to-one bundling must be no.
Bundling	Yes or no. Must be no if all-to-one bundling is yes and yes if all-to-one bundling is no.
All-to-one bundling	Yes or no. If yes, service multiplexing and bundling must be no. Must be no if bundling is yes.
Ingress and egress bandwidth profile per UNI	No or one of the following parameters: CIR, CBS, PIR, MBS.  If no, no bandwidth profile per UNI is set; otherwise, the traffic parameters CIR, CBS, PIR, and MBS need to be set.

*continues*

**Table 3-1** *UNI Service Attributes (Continued)*

<b>UNI Service Attribute</b>	<b>Parameter Values or Range of Values</b>
Ingress and egress bandwidth profile per EVC	No or one of the following parameters: CIR, CBS, PIR, MBS.
Ingress and egress bandwidth profile per CoS identifier	No or one of the following parameters: CIR, CBS, PIR, MBS.  If one of the parameters is chosen, specify the CoS identifier, Delay value, Jitter value, Loss value.  If no, no bandwidth profile per CoS identifier is set; otherwise, the traffic parameters CIR, CBS, PIR, and MBS need to be set.
Ingress and egress bandwidth profile per destination UNI per EVC	No or one of the following parameters: CIR, CBS, PIR, MBS.
Egress bandwidth profile per source UNI per EVC	No or one of the following parameters: CIR, CBS, PIR, MBS.
Layer 2 Control Protocol processing	Process, discard, or pass to EVC the following control protocol frames: <ul style="list-style-type: none"> <li>• IEEE 802.3x MAC control</li> <li>• Link Aggregation Control Protocol (LACP)</li> <li>• IEEE 802.1x port authentication</li> <li>• Generic Attribute Registration Protocol (GARP)</li> <li>• STP</li> <li>• Protocols multicast to all bridges in a bridged LAN</li> </ul>
UNI service activation time	Time value

**Table 3-2** *EVC Service Attributes*

<b>EVC Service Attribute</b>	<b>Type of Parameter Value</b>
EVC Type	P2P or MP2MP
CE-VLAN ID preservation	Yes or no
CE-VLAN CoS preservation	Yes or no
Unicast frame delivery	Discard, deliver unconditionally, or deliver conditionally for each ordered UNI pair
Multicast frame delivery	Discard, deliver unconditionally, or deliver conditionally for each ordered UNI pair
Broadcast frame delivery	Discard, deliver unconditionally, or deliver conditionally for each ordered UNI pair

**Table 3-2** *EVC Service Attributes (Continued)*

<b>EVC Service Attribute</b>	<b>Type of Parameter Value</b>
Layer 2 Control Protocol processing	Discard or tunnel the following control frames: <ul style="list-style-type: none"> <li>• IEEE 802.3x MAC control</li> <li>• Link Aggregation Control Protocol (LACP)</li> <li>• IEEE 802.1x port authentication</li> <li>• Generic Attribute Registration Protocol (GARP)</li> <li>• STP</li> <li>• Protocols multicast to all bridges in a bridged LAN</li> </ul>
EVC service activation time	Time value
EVC availability	Time value
EVC mean time to restore	Time value
Class of service	CoS identifier, Delay value, Jitter value, Loss value This assigns the Class of Service Identifier to the EVC

## Example of an L2 Metro Ethernet Service

This section gives an example of an L2 metro Ethernet service and how all the parameters defined by the MEF are applied. The example attempts to highlight many of the definitions and concepts discussed in this chapter.

If you have noticed, the concept of VPNs is inherent in L2 Ethernet switching. The carrier VLAN is actually a VPN, and all customer sites within the same carrier VLAN form their own user group and exchange traffic independent of other customers on separate VLANs.

The issue of security arises in dealing with VLAN isolation between customers; however, because the metro network is owned by a central entity (such as the metro carrier), security is enforced. First of all, the access switches in the customer basement are owned and administered by the carrier, so physical access is prevented. Second, the VLANs that are switched in the network are assigned by the carrier, so VLAN isolation is guaranteed. Of course, misconfiguration of switches and VLAN IDs could cause traffic to be mixed, but this problem can occur with any technology used, not just Ethernet. Issues of security always arise in public networks whether they are Ethernet, IP, MPLS, or Frame Relay networks. The only definite measure to ensure security is to have the customer-to-customer traffic encrypted at the customer sites and to have the customers administer that encryption.

Figure 3-9 shows an example of an L2 metro Ethernet VPN. This example attempts to show in a practical way how many of the parameters and the concepts that are discussed in this chapter are used.

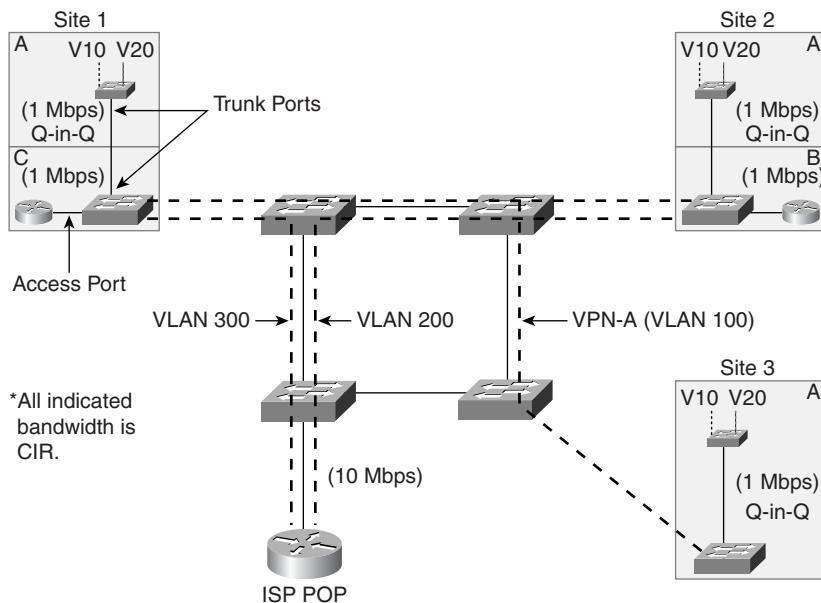
**Figure 3-9** All-Ethernet L2 Metro Service Example

Figure 3-9 shows a metro carrier offering an L2 MP2MP VPN service to customer A and a packet leased-line service (comparable to a traditional T1 leased line) to an ISP. In turn, the ISP is offering Internet service to customers B and C. It is assumed that customer A connects to the carrier via L2 Ethernet switches and customers B and C connect via IP routers. Notice the difference between access ports and trunk ports on the Ethernet switches. The ports connecting the customer's Ethernet switch to the carrier's Ethernet switch are trunk ports, because these ports are carrying multiple VLANs between the two switches. When the carrier's switch port is configured for Q-in-Q, it encapsulates the customers' CE-VLAN tags VLAN 10 and VLAN 20 inside the carrier VLAN 100. On the other hand, the ports connecting the customer with the carrier switch are access ports and are carrying untagged traffic from the router. Tables 3-3 and 3-4 describe the UNI and EVC service attributes for customers A, B, and C as defined by the MEF.

**Table 3-3** Customer A E-LAN UNI Service Attributes

Customer A E-LAN UNI Service Attribute	Parameter Values or Range of Values
Physical medium	Standard Ethernet physical interfaces
Speed	100 Mbps site 1, 10 Mbps sites 2 and 3
Mode	Full duplex all sites
MAC layer	IEEE 802.3-2002
Service multiplexing	No

**Table 3-3** *Customer A E-LAN UNI Service Attributes (Continued)*

<b>Customer A E-LAN UNI Service Attribute</b>	<b>Parameter Values or Range of Values</b>
Bundling	No
All-to-one bundling	Yes
Ingress and egress bandwidth profile per CoS identifier	<p>All sites CoS 1:</p> <ul style="list-style-type: none"> <li>• CIR = 1 Mbps, CBS = 100 KB, PIR = 2 Mbps, MBS = 100 KB</li> <li>• CoS ID = 802.1p 6–7</li> <li>• Delay &lt; 10 ms, Loss &lt; 1%</li> </ul> <p>All sites CoS 0:</p> <ul style="list-style-type: none"> <li>• CIR = 1 Mbps, CBS = 100 KB, PIR = 10 Mbps, MBS = 100 KB</li> <li>• CoS ID = 802.1p 0–5, Delay &lt; 35 ms, Loss &lt; 2%</li> </ul>
Layer 2 Control Protocol processing	<ul style="list-style-type: none"> <li>• Process IEEE 802.3x MAC control</li> <li>• Process Link Aggregation Control Protocol (LACP)</li> <li>• Process IEEE 802.1x port authentication</li> <li>• Pass Generic Attribute Registration Protocol (GARP)</li> <li>• Pass STP</li> <li>• Pass protocols multicast to all bridges in a bridged LAN</li> </ul>
UNI service activation time	One hour after equipment is installed

Note in Table 3-3 that customer A is given only one MP2P EVC; hence, there is no service multiplexing. All customer VLANs 10 and 20 are mapped to the MP2MP EVC in the form of carrier VLAN 100. Customer A is given two Class of Service profiles—CoS 1 and CoS 0. Each profile has its set of performance attributes. Profile 1, for example, is applied to high-priority traffic, as indicated by 802.1p priority levels 6 and 7. Profile 0 is lower priority, with less-stringent performance parameters. For customer A, the metro carrier processes the 802.3x and LACP frames on the UNI connection and passes other L2 control traffic that belongs to the customer. Passing the STP control packets, for example, prevents any potential loops within the customer network, in case the customer has any L2 backdoor direct connection between its different sites.

**Table 3-4** *Customer A E-LAN EVC Service Attributes*

<b>Customer A E-LAN EVC Service Attribute</b>	<b>Type of Parameter Value</b>
EVC type	MP2MP
CE-VLAN ID preservation	Yes

**Table 3-4** *Customer A E-LAN EVC Service Attributes (Continued)*

<b>Customer A E-LAN EVC Service Attribute</b>	<b>Type of Parameter Value</b>
CE-VLAN CoS preservation	Yes
Unicast frame delivery	Deliver unconditionally for each UNI pair
Multicast frame delivery	Deliver unconditionally for each UNI pair
Broadcast frame delivery	Deliver unconditionally for each UNI pair
Layer 2 Control Protocol processing	Tunnel the following control frames: <ul style="list-style-type: none"> <li>• IEEE 802.3x MAC control</li> <li>• Link Aggregation Control Protocol (LACP)</li> <li>• IEEE 802.1x port authentication</li> <li>• Generic Attribute Registration Protocol (GARP)</li> <li>• STP</li> <li>• Protocols multicast to all bridges in a bridged LAN</li> </ul>
EVC service activation time	Twenty minutes after UNI is operational
EVC availability	Three hours
EVC mean time to restore	One hour
Class of service	All sites CoS 1: <ul style="list-style-type: none"> <li>• CoS ID = 802.1p 6–7</li> <li>• Delay &lt; 10 ms, Loss &lt; 1%, Jitter (value)</li> </ul> All sites CoS 0: <ul style="list-style-type: none"> <li>• CoS ID = 802.1p 0–5, Delay &lt; 35 ms, Loss &lt; 2%, Jitter (value)</li> </ul>

The EVC service parameters for customer A indicate that the EVC is an MP2MP connection and the carrier transparently moves the customer VLANs between sites. The carrier does this using Q-in-Q tag stacking with a carrier VLAN ID of 100. The carrier also makes sure that the 802.1p priority fields that the customer sends are still carried within the network. Note that the carrier allocates priority within its network whichever way it wants as long as the carrier delivers the SLA agreed upon with the customer as described in the CoS profiles. For customer A, the carrier passes all unicast, multicast, and broadcast traffic and also tunnels all L2 protocols between the different sites.

Tables 3-5 and 3-6 describe customers B and C and ISP POP service profile for the Internet connectivity service. These are the service attributes and associated parameters for customers



B and C as well as the service attributes and associated parameters for the ISP POP offering Internet connectivity to these customers.

**Table 3-5** *Customers B and C and ISP POP UNI Service Attributes*

<b>Customers B and C and ISP POP Internet Access UNI Service Attribute</b>	<b>Parameter Values or Range of Values</b>
Physical medium	Standard Ethernet physical interfaces
Speed	10 Mbps for customers B and C, 100 Mbps for the ISP POP
Mode	Full duplex all sites
MAC layer	IEEE 802.3-2002
Service multiplexing	Yes, only at ISP POP UNI
Bundling	No
All-to-one bundling	No
Ingress and egress bandwidth profile per EVC	<p>Customers B and C</p> <p>CIR = 1 Mbps, CBS = 100 KB, PIR = 2 Mbps, MBS = 100 KB</p> <p>ISP POP</p> <p>CIR = 10 Mbps, CBS = 1 MB, PIR = 100 Mbps, MBS = 1 MB</p>
Layer 2 Control Protocol processing	<p>Discard the following control frames:</p> <ul style="list-style-type: none"> <li>• IEEE 802.3x MAC control</li> <li>• Link Aggregation Control Protocol (LACP)</li> <li>• IEEE 802.1x port authentication</li> <li>• Generic Attribute Registration Protocol (GARP)</li> <li>• STP</li> <li>• Protocols multicast to all bridges in a bridged LAN</li> </ul>
UNI service activation time	One hour after equipment is installed

For customers B and C and ISP POP UNI service parameters, because two different P2P EVCs (carrier VLANs 200 and 300) are configured between the customers and the ISP POP, service multiplexing occurs at the ISP UNI connection where two EVCs are multiplexed on the same physical connection. For this Internet access scenario, routers are the customer premises equipment, so it is unlikely that the customer will send any L2 control-protocol packets to the carrier. In any case, all L2 control-protocol packets are discarded if any occur.

**Table 3-6** *Customers B and C and ISP POP EVC Service Attributes*

<b>Customers B and C and ISP POP Internet Access EVC Service Attribute</b>	<b>Type of Parameter Value</b>
EVC type	P2P
CE-VLAN ID preservation	No; mapped VLAN ID for provider use
CE-VLAN CoS preservation	No
Unicast frame delivery	Deliver unconditionally for each UNI pair
Multicast frame delivery	Deliver unconditionally for each UNI pair
Broadcast frame delivery	Deliver unconditionally for each UNI pair
Layer 2 Control Protocol processing	Discard the following control frames: <ul style="list-style-type: none"> <li>• IEEE 802.3x MAC control</li> <li>• Link Aggregation Control Protocol (LACP)</li> <li>• IEEE 802.1x port authentication</li> <li>• Generic Attribute Registration Protocol (GARP)</li> <li>• STP</li> <li>• Protocols multicast to all bridges in a bridged LAN</li> </ul>
EVC service activation time	Twenty minutes after UNI is operational
EVC availability	Three hours
EVC mean time to restore	One hour
Class of service	One CoS service is supported: Delay < 30 ms, Loss < 1%, Jitter (value)

The EVC parameters indicate that the carrier is not preserving any customer VLANs or CoS info. Also, because this is an Internet access service, normally the provider receives untagged frames from the CPE router. The provider can map those frames to carrier VLANs 200 and 300 if it needs to separate the traffic in its network. The VLAN IDs are normally stripped off before given to the ISP router.

## Challenges with All-Ethernet Metro Networks

All-Ethernet metro networks pose many scalability and reliability challenges. The following are some of the issues that arise with an all-Ethernet control plane:

- Restrictions on the number of customers
- Service monitoring
- Scaling the L2 backbone

- Service provisioning
- Interworking with legacy deployments

The following sections describe each of these challenges.

## Restrictions on the Number of Customers

The Ethernet control plane restricts the carrier to 4096 customers, because the 802.1Q defines 12 bits that can be used as a VLAN ID, which restricts the number of VLANs to  $2^{12} = 4096$ . Remember that although Q-in-Q allows the customer VLANs (CE-VLANs) to be hidden from the carrier network, the carrier is still restricted to 4096 VLAN IDs that are global within its network. For many operators that are experimenting with the metro Ethernet service, the 4096 number seems good enough for an experimental network but presents a long-term roadblock if the service is to grow substantially.

## Service Monitoring

Ethernet does not have an embedded mechanism that lends to service monitoring. With Frame Relay LMI, for example, service monitoring and service integrity are facilitated via messages that report the status of the PVC. Ethernet service monitoring requires additional control-plane intelligence. New Link Management Interface (LMI) protocols need to be defined and instituted between the service provider network and the CPE to allow the customer to discover the different EVCs that exist on the UNI connection. The LMI could learn the CE-VLAN to EVC map and could learn the different service parameters such as bandwidth profiles. Other protocols need to be defined to discover the integrity of the EVC in case of possible failures. You have seen in the previous section how performance parameters could indicate the availability of an EVC. Protocols to extract information from the UNI and EVC are needed to make such information usable.

## Scaling the L2 Backbone

A metro carrier that is building an all-Ethernet network is at the mercy of STP. STP blocks Ethernet ports to prevent network loops. Traffic engineering (discussed in Chapter 5, “MPLS Traffic Engineering”) is normally a major requirement for carriers to have control over network bandwidth and traffic trajectory. It would seem very odd for any carrier to have the traffic flow in its network be dependant on loop prevention rather than true bandwidth-optimization metrics.

## Service Provisioning

Carrying a VLAN through the network is not a simple task. Any time a new carrier VLAN is created (a new VPN), care must be taken to configure that VLAN across all switches that need to participate in that VPN. The lack of any signaling protocols that allow VPN information to

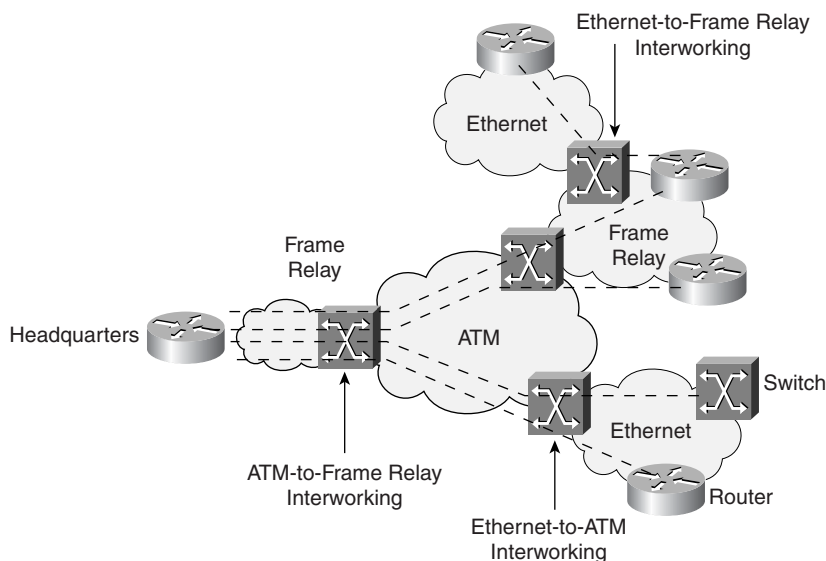
be exchanged makes the task manual and tedious. Early adopters of metro Ethernet have endured the pains of carrying VLANs across many switches. Even with the adoption of new protocols such as 802.1s (“Amendment to 802.1Q (TM) Virtual Bridged Local Area Networks: Multiple Spanning Trees”), the task of scaling the network is almost impossible.

## Interworking with Legacy Deployments

Another challenge facing Ethernet deployments is interworking with legacy deployments such as existing Frame Relay and ATM networks. Frame Relay has been widely deployed by many enterprises as a WAN service. Remote offices are connected to headquarters via P2P Frame Relay circuits forming a hub-and-spoke topology. Enterprises that want to adopt Ethernet as an access technology expect the carrier to provide a means to connect the new sites enabled with Ethernet access with existing headquarters sites already enabled with Frame Relay. This means that a function must exist in the network that enables Frame Relay and Ethernet services to work together.

The IETF has standardized in RFC 2427, *Multiprotocol Interconnect over Frame Relay*, how to carry different protocols over Frame Relay, including Ethernet. In some other cases, the Ethernet and Frame Relay access networks are connected by an ATM core network. In this case, two service-interworking functions need to happen, one between Ethernet and ATM and another between ATM and Frame Relay. Ethernet-to-ATM interworking is achieved using RFC 2684, and ATM-to-Frame Relay interworking is achieved via the Frame Relay Forum specification FRF 8.1. Figure 3-10 illustrates the service-interworking functions.

**Figure 3-10** *Service Interworking*



JUNIPER Exhibit 1003



App. 6, pg. 89

Figure 3-10 shows a scenario in which an enterprise headquarters is connected to its remote sites via Frame Relay connections carried over an ATM network. The different service-interworking functions are displayed to allow such networks to operate. For service interworking, two encapsulation methods are defined: one is bridged, and the other is routed. Both sides of the connection are either bridged or routed. Some challenges might exist if one end of the connection is connected to a LAN switch, and hence bridged, while the other end is connected to a router. Other issues will arise because of the different Address Resolution Protocol (ARP) formats between the different technologies, such as Ethernet, Frame Relay, and ATM. Some vendors are attempting to solve these problems with special software enhancements; however, such practices are still experimental and evolving.

It is all these challenges that motivated the emergence of hybrid architectures consisting of multiple L2 domains that are connected via L3 IP/MPLS cores. The network can scale because L2 Ethernet would be constrained to more-controlled access deployments that limit the VLAN and STP inefficiencies. The network can then be scaled by building a reliable IP/MPLS core. This is discussed in Chapter 4, “Hybrid L2 and L3 IP/MPLS Networks.”

## Conclusion

This chapter has discussed many aspects of metro Ethernet services. The MEF is active in defining the characteristics of these services, including the service definitions and framework and the many service attributes that make up the services. Defining the right traffic and performance parameters, class of service, service frame delivery, and other aspects ensures that buyers and users of the service understand what they are paying for and also helps service providers communicate their capabilities.



This chapter covers the following topics:

- Understanding VPN Components
- Delivering L3VPNs over IP
- L2 Ethernet Services over an IP/MPLS Network

# Hybrid L2 and L3 IP/MPLS Networks

In Chapter 3, “Metro Ethernet Services,” you reviewed the issues that can be created by an L2-only Ethernet model. This chapter first focuses on describing a pure L3VPN implementation and its applicability to metro Ethernet. This gives you enough information to compare L3VPNs and L2VPNs relative to metro Ethernet applications. The chapter then delves into the topics of deploying L2 Ethernet services over a hybrid L2 Ethernet and L3 IP/MPLS network. Some of the basic scalability issues to be considered include restrictions on the number of customers because of the VLAN-ID limitations, scaling the L2 backbone with spanning tree, service provisioning and monitoring, and carrying VLAN information within the network. The following section describes some basic VPN definitions and terminology.

## Understanding VPN Components

There are normally two types of VPNs: customer premises equipment-(CPE) based VPNs and network-based VPNs. With CPE-based VPNs, secure connections are created between the different customer premises equipment to form a closed user group/VPN. This normally creates scalability issues, because many CPE devices have to be interconnected in a full mesh or a partial mesh to allow point-to-multipoint connectivity. On the other hand, network-based VPNs create some level of hierarchy where connections from many CEs are aggregated into an edge switch or router offering the VPN service.

The definitions of the different elements of the network follow:

- **Customer edge (CE)**—The customer edge device resides at the edge of the enterprise. This device is usually a router or a host in L3VPNs; however, as you will see with L2VPNs, the CE could also be an L2 switch. The CE connects to the provider network via different data-link protocols such as PPP, ATM, Frame Relay, Ethernet, GRE, and so on.
- **Provider edge (PE)**—The provider edge device is a provider-owned device that offers the first level of aggregation for the different CEs. The PE logically separates the different VPNs it participates in. The PE does not have to participate in all VPNs but would only participate in the VPNs of the enterprises that are directly attached to it.

- **Provider (P)**—The provider device is normally a core IP/MPLS router that offers a second level of aggregation for the PEs. This device does not participate in any VPN functionality and is normally agnostic to the presence of any VPNs.

The remainder of the chapter mainly focuses on different types of VPNs and how they differ between an L2 or L3 service. The different types of VPNs include

- GRE- and MPLS-based L3VPNs
- Hybrid Ethernet and IP/MPLS L2VPNs via L2TPv3, Ethernet over MPLS (EoMPLS), and Virtual Private LAN Service (VPLS)

## Delivering L3VPNs over IP

L3VPNs allow the provider to extend its customer's private IP network over the provider's backbone. When delivering an L3 service, the service provider is normally involved in the assignment and management of a pool of IP addresses allocated to its customer. This is typical of carriers that are also ISPs offering Internet services or carriers offering IP multicast services and so on. L3VPNs can be delivered via GRE tunnels or MPLS L3VPNs.

### GRE-Based VPNs

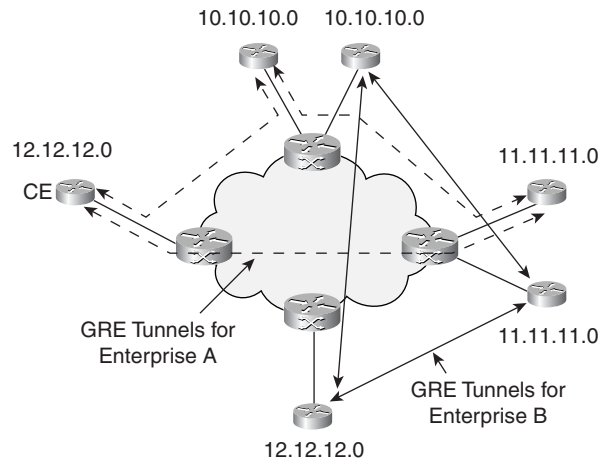
L3VPN services over IP have traditionally been done using generic routing encapsulation (GRE) tunnels, which allow the encapsulation of IP packets inside IP packets. GRE-based VPNs are CE-based VPNs. A network hierarchy can be maintained in which an enterprise that has, for example, a private IP addressing scheme can create a private VPN on top of a service provider's network. IP forwarding is used to exchange traffic between the endpoints of GRE tunnels, allowing full or partial connectivity between the different sites of the same enterprise. From a scalability perspective, this scheme could scale to a certain point and then become unmanageable, because the VPN becomes the collection of many point-to-point tunnels. As many sites are added to the VPN and many tunnels have to be created to all or a partial set of the other sites, the operational management of such a scheme becomes cost-prohibitive, especially because there are no rules or guidelines or an industry push to allow such tunneling schemes to scale.

Figure 4-1 shows an example of a service provider delivering a GRE-based VPN service using managed CEs located at different enterprise sites. The provider is managing the CEs at each site of each enterprise and is managing the tunnel connectivity between the different sites. As the number of enterprises grows and the number of sites per enterprise grows as well, this model will definitely have scalability issues. Notice that different enterprises could use overlapping private IP addresses, because all IP and routing information between the enterprise sites is carried within tunnels and hence is hidden from the provider's network and other enterprises.

For large-scale deployments of IP VPNs, the industry has gradually moved toward adopting MPLS L3VPNs, as defined in RFC 2547.



Figure 4-1 GRE Tunnels



## MPLS L3VPNs

MPLS L3VPNs are network-based VPNs. This scheme defines a scalable way for the service provider to offer VPN services for enterprises. Enterprises can leverage the service provider backbone to globally expand their intranets and extranets. An *intranet* normally means that all sites in the VPN connect to the same customer, and *extranet* means that the various sites in the VPN are owned by different enterprises, such as the suppliers of an enterprise. An example of an extranet would be a car manufacturer that builds a network that connects it and all its parts suppliers in a private network.

Although MPLS L3VPNs provide a sound and scalable solution for delivering VPNs over IP, they have some characteristics that make them overkill for metro Ethernet services. L3VPNs, for example, are more adequate for delivering IP services than L2VPN services. This is one of the reasons that the industry is looking at L2VPNs for metro Ethernet services. To understand the differences between L2VPNs and L3VPNs, it helps to identify the different elements of MPLS L3VPNs (RFC 2547) and the challenges that come with them.

MPLS L3VPNs use the CE, PE, and P terminology described earlier in the “Understanding VPN Components” section. In the case where the CE is a router, the CE and PE become routing peers if a routing protocol is used between the two to exchange IP prefixes. In other scenarios, static routing is used between the PE and CE to alleviate the exchange of routing information. With L3VPNs, enterprise edge routers have to talk only to their direct neighbor, which is the router owned by the provider. From a scalability perspective, the L3VPN model scales very well, because each site does not need to know of the existence of other sites. On the other hand, this model is not so good for enterprises that would like to maintain their own internal routing practices and control the routing mechanism used between the different sites. Also, this model forces the service provider to participate in and manage the IP addressing schemes for its customers, as is typically done when IP services are sold. This model is not

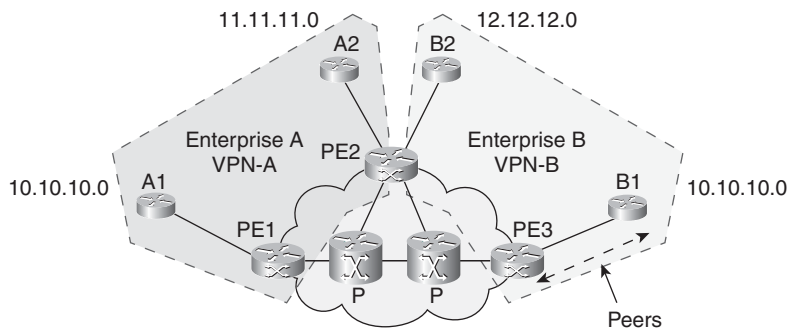
adequate for selling L2 services only (L2VPN) where the customer's IP network becomes an overlay on top of the service provider's network.

Another disadvantage of L3VPNs when used for metro Ethernet services is that L3VPNs apply only to the transport of IPv4 packets. For metro deployments, enterprise traffic consists of IPv4 as well as other types of traffic such as IPX and SNA. An L2VPN allows any type of traffic to be encapsulated and transported across the metro network.

Take a close look at the example in Figure 4-2. The provider is delivering VPN services to two different enterprises, A and B, and each enterprise has two different sites. Sites A1 and A2 are part of enterprise A and belong to VPN-A. Sites B1 and B2 are part of enterprise B and belong to VPN-B. Note that enterprises A and B could have overlapping IP addresses. The following are the reasons why the MPLS L3VPN model scales:

- Each PE knows only of the VPNs it attaches to. PE1 knows only of VPN-A, and PE3 knows only of VPN-B.
- The P routers do not have any VPN information.
- The CE routers peer with their directly attached PEs. A1 peers with PE1, B1 peers with PE3, and so on.

**Figure 4-2** *MPLS L3VPN Principles*



The following sections describe

- How MPLS L3VPN PEs maintain separate forwarding tables between different VPNs
- The concept of VPN-IPv4 addresses
- How packets are transported across the backbone using the MPLS L3VPN mechanism

## Maintaining Site Virtual Router Forwarding Tables

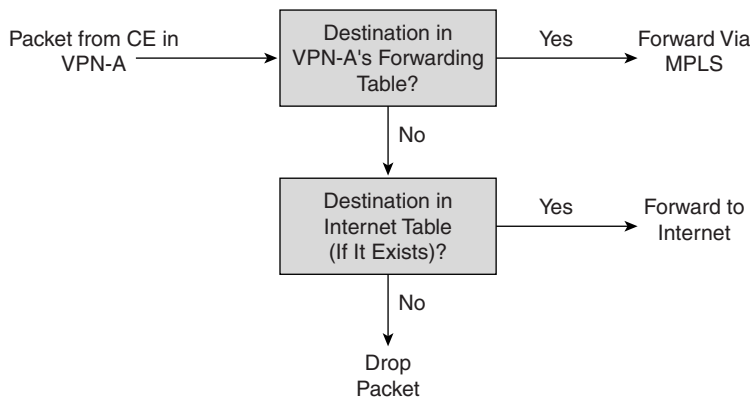
The fundamental operation of the MPLS L3VPN model follows:

- Each PE router maintains a separate virtual router forwarding (VRF) table for each site the PE is attached to. The forwarding table contains the routes to all other sites that participate in a set of VPNs.

- The PEs populate the forwarding tables from information learned from the directly attached sites or learned across the backbone from other PEs that have a VPN in common. Information from directly attached CEs is learned via routing protocols such as OSPF, IS-IS, RIP, and BGP or via static configuration. Distribution of VPN information across the backbone is done via *multiprotocol BGP (MP-BGP)*. MP-BGP introduces extensions to the BGP-4 protocol to allow IPv4 prefixes that are learned from different VPNs to be exchanged across the backbone. IP prefixes can overlap between different VPNs via the use of VPN-IPv4 address, as explained later, in the section “Using VPN-IPv4 Addresses in MPLS L3VPNs.”
- The CEs learn from the PEs about the routes they can reach via routing protocols or static configuration.

Traffic is forwarded across the backbone using MPLS. MPLS is used because the backbone P routers have no VPN routes; hence, traditional IP routing cannot be used. Figure 4-3 illustrates the packet forwarding process.

**Figure 4-3** *The Packet Forwarding Process*

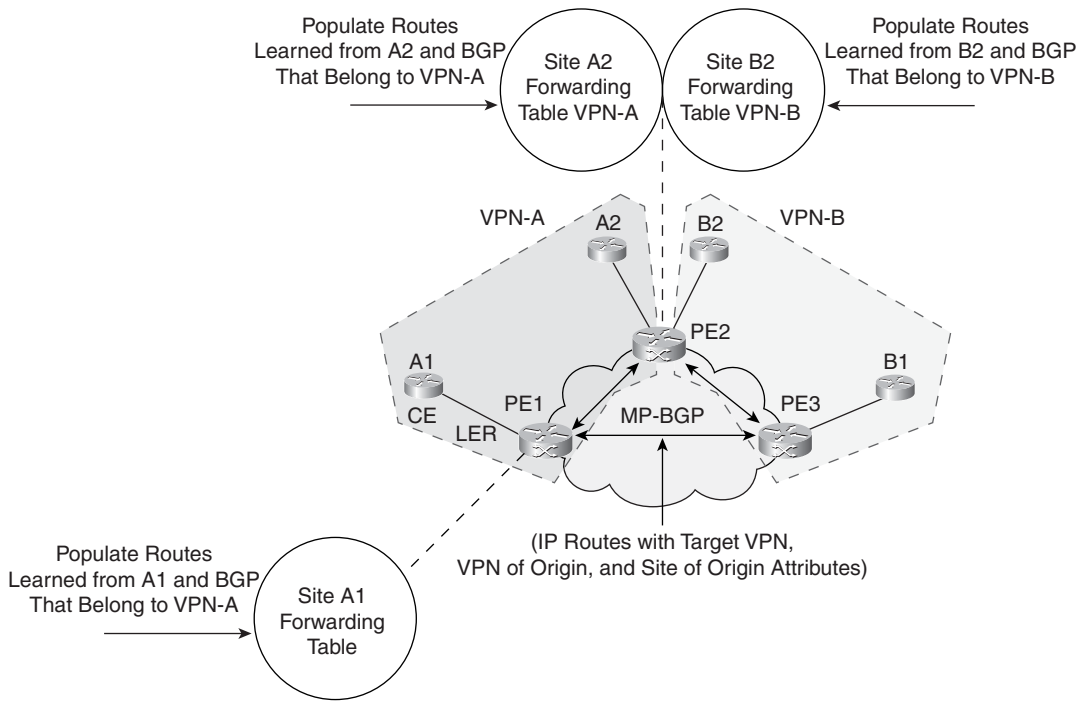


As Figure 4-3 shows, when a packet is received from a site, the PE looks up the IP destination in the site’s forwarding table and, if found, forwards the packet via MPLS. Otherwise, the PE checks the destination in other forwarding tables and discards the packet if no match has been made.

Figure 4-4 shows how the PEs maintain a different forwarding table per site. PE1 contains a forwarding table for enterprise A site 1 (A1). That forwarding table is populated from routes learned from A1 and from BGP routes across the backbone. PE2 contains a forwarding table for both enterprise A site 2 (A2) and enterprise B site 2 (B2).

MPLS L3VPNs use target VPNs, VPN of origin, and site of origin to be able to identify and separate the different routes belonging to different VPNs and to clearly identify the origin of a particular route. The following sections describe these features.

Figure 4-4 PE Logical Separation



### Target VPN

For identifying different VPNs, every per-site forwarding table is associated with one or more target VPN attributes. When a PE router creates a VPN-IPv4 route, the route is associated with one or more target VPN attributes, which are carried in BGP as attributes of the route. A *route attribute* is a parameter that gives the route special characteristics and is a field that is distributed inside a BGP advertisement. The target VPN attribute identifies a set of sites. Associating a particular target VPN attribute with a route allows the route to be placed in the per-site forwarding tables used for routing traffic that is received from the corresponding site. In Figure 4-4, when PE1 receives BGP routes from PE2, PE1 installs in the A1 forwarding table only routes that have a target VPN attribute VPN-A. This ensures that PE1 does not contain any routes to VPN-B, because PE1 does not have any attached sites that belong to VPN-B. On the other hand, PE2 installs in the respective forwarding tables routes that belong to VPN-A and VPN-B, because PE2 is attached to sites that belong to both VPNs.

**NOTE**

In the context of MPLS L3VPN, IPv4 addresses are referred to as VPN-IPv4 addresses. The section “Using VPN-IPv4 Addresses in MPLS L3VPNs” discusses scenarios for VPN-IPv4 addresses in more detail.

## VPN of Origin

Additionally, a VPN-IPv4 route may be optionally associated with a VPN of origin attribute. This attribute uniquely identifies a set of sites and identifies the corresponding route as having come from one of the sites in that set. Typical uses of this attribute might be to identify the enterprise that owns the site where the route leads, or to identify the site's intranet. However, other uses are also possible, such as to identify which routes to accept and which to drop based on the VPN of origin. By using both the target VPN and the VPN of origin attributes, different kinds of VPNs can be constructed.

## Site of Origin

Another attribute, called the site of origin attribute, uniquely identifies the site from which the PE router learned the route (this attribute could be encoded as an extended BGP community attribute). All routes learned from a particular site must be assigned the same site of origin attribute.

## Using VPN-IPv4 Addresses in MPLS L3VPNs

The purpose of VPN-IPv4 addresses is to allow routers to create different routes to a common IPv4 address. This is useful in different scenarios that relate to L3VPNs.

One such scenario occurs when multiple VPNs have overlapping IPv4 addresses. In this case, routers need to treat each address differently when populating the per-site forwarding table. If the same address belongs in two different VPNs, the router needs to place the same address into two different VRF tables. Another use of VPN-IPv4 is to create separate routes to reach the same IPv4 destination address. In the case of an enterprise that has an intranet and an extranet, the same server can have its IP address advertised in two different routes, one used by the intranet and another by the extranet. The extranet route could be forced to go through a firewall before reaching the server.

The VPN-IPv4 address, as shown in Figure 4-5, is a 12-byte quantity, beginning with an 8-byte Route Distinguisher (RD) and ending with a 4-byte IPv4 address. The RD consists of a Type field that indicates the length of the Administrator and Assigned Number fields. The Administrator field identifies an Assigned Number authority field, such as an autonomous system number given to a certain service provider. The service provider can then allocate the assigned number to be used for a particular purpose. Note that the RD by itself does not contain enough information to indicate the origin of the route or to which VPNs the route needs to be distributed. In other words, the RD is not indicative of a particular VPN. The purpose of the RD is only to allow the router to create different routes to the same IPv4 address.

As referenced in Figure 4-5, ISP A wants to distinguish between two IPv4 addresses 10.10.10.0; therefore, it assigns these two addresses two different RDs. The RD administrator number is ISP A's autonomous system (AS) number (1111). The assigned numbers 1 and 2 are just arbitrary numbers that help the routers distinguish between the two IP addresses that could be in the same VPN or different VPNs. Again, it is important to understand that the VPN-IPv4 does not modify

the IP address itself but rather is an attribute sent within BGP (RFC 2283) that indicates that the IP address belongs to a certain family.

**Figure 4-5** *VPN-IPv4 Address*

Route Distinguisher (RD)			
Type (2 Bytes)	Administrator (2 Bytes)	Assigned Number (4 Bytes)	IPv4 (4 Bytes)
	AS 1111 (ISP A)	1	10.10.10.0
	AS 1111 (ISP A)	2	10.10.10.0

### Forwarding Traffic Across the Backbone

Only the edge PE routers have information about the VPN IP prefixes. The backbone P routers do not carry any VPN IP prefixes. With traditional IP forwarding, this model does not work, because the P routers drop any traffic destined for the VPN IP addresses. MPLS is used to allow packet forwarding based on labels rather than IP addresses. The PE routers tag the traffic with the right label based on the destination IP address it needs to go to, and the MPLS P routers switch the traffic based on the MPLS labels. If this model is not adopted, the P routers would have to carry IP prefixes for all VPNs, which would not scale.

MPLS L3VPN does not mandate the use of traffic engineering (a topic that is explained in more detail in Chapter 5, “MPLS Traffic Engineering”). When traffic moves from one site to another across the carrier’s backbone, it follows the MPLS label switched path (LSP) assigned for that traffic. The LSP itself could have been formed via dynamic routing calculated by the routing protocols. On the other hand, the LSP could be traffic-engineered to allow certain types of traffic to follow a well-defined trajectory. Also, many mechanisms can be used for traffic rerouting in case of failure. The mechanism used depends on whether the carrier requires normal IP routing or MPLS fast reroute mechanisms (as explained in Chapter 6, “RSVP for Traffic Engineering and Fast Reroute”).

The traffic across the MPLS backbone carries a label stack. The label on top of the stack is called the packet-switched network (PSN) tunnel label and is indicative of the path that a packet needs to take from the ingress PE to the egress PE. The label beneath is indicative of the particular VPN that the packet belongs to. In the case where IP forwarding (rather than MPLS) is used in the provider routers, the PSN tunnel can be replaced by a GRE tunnel, and the packet would carry the VPN label inside the GRE tunnel.

### Applicability of MPLS L3VPNs for Metro Ethernet

The MPLS L3VPN model presents many challenges if used to deliver metro Ethernet services. This model is more applicable for delivering IP services, where an enterprise is outsourcing the

operation of its WAN/metro IP network to a service provider. From an administration point of view, the MPLS L3VPN model dictates that the carrier is involved with the customer's IP addressing scheme. Remember that the CE routers would have to peer with the provider's routers. If static routing is not used, routing exchange between the PE and CE might involve configuring routing protocols like RIP and OSPF and will involve many guidelines to allow protocols such as OSPF to understand the separation between the different VPN routes and to distribute the correct routes between BGP and OSPF.

From an equipment vendor perspective, while the MPLS L3VPN model scales in theory, it introduces major overhead on the edge routers. If you assume that an edge router needs to support 1000 VPNs, and each VPN has 1000 IP prefixes, the edge routers would have to maintain at least 1,000,000 IP prefixes in 1000 separate forwarding tables. Most routers on the market are still struggling to reach 256,000 to 500,000 IP entries, depending on the vendor's implementation. So what would happen if the IP prefixes per VPN reaches 5000 entries rather than 1000? A clever answer would be to support 200 VPNs per PE router to stay within the 1,000,000 prefix limit, until vendors find a way to increase that number.

There are other ways of deploying L3VPNs, such as using virtual routers where different instances of routing protocols run on each router. Each routing instance carries the IP prefixes of a different VPN, and traffic is forwarded across the network using traditional IP forwarding; hence, the final outcome is very similar to running MPLS L3VPNs. The MPLS L3VPN and virtual routers have their advantages in delivering IP services, which include IP QoS mechanisms, IP address pool management, and so on. These advantages are very important but are outside the scope of this book and will not be discussed. However, L3VPN is still overkill for deploying metro Ethernet services, which focuses on simpler deployments and L2 services.

It is understandable why the industry started looking at simpler VPN schemes like L2VPNs to avoid many of the L3VPN complexities and to create a model in which simpler services like Transparent LAN Service (TLS) can be deployed with less operational overhead.

In an L2 service, the carrier offers its customers the ability to "transparently" overlay their own networks on top of the carrier's network. The customer of a carrier could be an ISP that offers Internet services and purchases last-mile connectivity from the carrier, or an enterprise customer that uses the carrier's backbone to build the enterprise WAN while still controlling its internal IP routing.

## L2 Ethernet Services over an IP/MPLS Network

The inherent properties of an IP/MPLS network mitigate most of the scalability issues by design. IP and MPLS have been widely deployed in large service provider networks, and these protocols have been fine-tuned over the years to offer high levels of stability and flexibility. Table 4-1 shows a brief comparison of the merits of IP/MPLS versus L2 Ethernet networks.

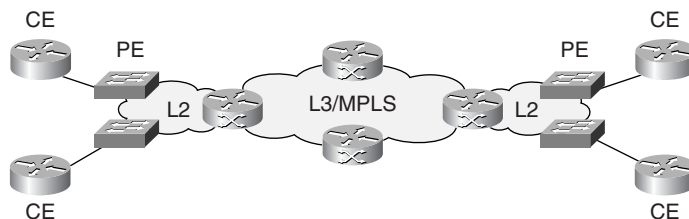
**Table 4-1** *Comparing Ethernet and IP/MPLS*

Feature	Ethernet	IP/MPLS
Signaling	No signaling	LDP, RSVP-TE, and so on
Loop-free topology	Blocked ports via Spanning Tree Protocol	Yes, via routing protocols and Time To Live
User and service identification	VLAN ID space limited	Label space more scalable
Traffic engineering (TE)	No TE	RSVP-TE
Restoration	Via STP	Backup path, MPLS fast reroute
Address aggregation	No aggregation for MAC addresses	Yes, via classless interdomain routing

Segmenting the L2 Ethernet network with IP/MPLS creates an L2 Ethernet domain at the metro access and an IP/MPLS metro edge/core and WAN backbone capable of carrying the L2 services. As you will see in this chapter, the closer the IP/MPLS network gets to the customer, the more scalable the service becomes; however, it introduces more complications.

The exercise of deploying Ethernet L2 services becomes one of balance between the L2 Ethernet simplicity and its scalability shortfalls and the L3 IP/MPLS scalability and its complexity shortfalls. First, it helps to compare and contrast the benefits that IP/MPLS offers over flat L2 networks.

You have seen so far in this book two extremes: one with an MPLS L3VPN service and one with an all-L2 Ethernet service. In this chapter, you see the hybrid model that falls in between. Figure 4-6 shows how an IP/MPLS domain can create a level of hierarchy that allows the L2 services to be confined to the access/edge network. There could be either an L2 access with IP/MPLS edge and core or an L2 access and edge with IP/MPLS core.

**Figure 4-6** *Hybrid L2 and IP/MPLS Metro*

The IP/MPLS edge/core network limits the L2 domains to the access or access/edge side and provides a scalable vehicle to carry the L2 services across.

The L2 Ethernet service across an IP/MPLS cloud can be a point-to-point (P2P) or multipoint-to-multipoint (MP2MP) service. This is very similar to the Metro Ethernet Forum (MEF) definitions of an Ethernet Line Service (ELS) and Ethernet LAN Service (E-LAN). The following

**JUNIPER Exhibit 1003**

App. 6, pg. 101



associates the service with the different methods to deliver it:

- **P2P Ethernet Service**—Comparable to ELS, delivered via:
  - L2TPv3 over an IP network
  - Ethernet over MPLS, also known as draft-martini in reference to the author of the original draft
- **MP2MP Ethernet Service**—Comparable to E-LAN, delivered via VPLS

Before getting into more details of the different mechanisms to deploy P2P and MP2MP L2 services, it helps to understand the packet leased-line concept, which is also referred to as pseudowire (PW), as explained next.

## The Pseudowire Concept

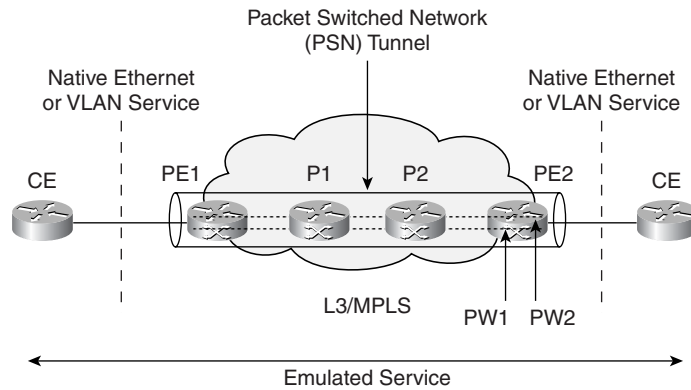
The Internet Engineering Task Force (IETF) has defined the concept of a pseudowire. An Ethernet PW allows Ethernet/802.3 protocol data units (PDUs) to be carried over a PSN, such as an IP/MPLS network. This allows service providers or enterprise networks to leverage an existing IP/MPLS network to offer Ethernet services.

You could set up the PW via manual configuration or a signaling protocol such as BGP or LDP. The PW may operate over an MPLS, IPv4, or IPv6 PSN.

An Ethernet PW emulates a single Ethernet link between exactly two endpoints. The PW terminates a logical port within the PE. This port provides an Ethernet MAC service that delivers each Ethernet packet that is received at the logical port to the logical port in the corresponding PE at the other end of the PW. Before a packet is inserted into the PW at the PE, the packet can go through packet processing functions that may include the following:

- Stripping
- Tag stacking or swapping
- Bridging
- L2 encapsulation
- Policing
- Shaping

Figure 4-7 shows a reference model that the IETF has adopted to support the Ethernet PW emulated services. As Figure 4-7 shows, multiple PWs can be carried across the network inside a bigger tunnel called the PSN tunnel. The PSN tunnel is a measure to aggregate multiple PWs into a single tunnel across the network. The PSN tunnel could be formed using generic routing encapsulation (GRE), Layer 2 Tunneling Protocol (L2TP), or MPLS and is a way to shield the internals of the network, such as the P routers, from information relating to the service provided by the PEs. In Figure 4-7, while the PE routers are involved in creating the PWs and mapping the L2 service to the PW, the P routers are agnostic to the L2 service and are passing either IP or MPLS packets from one edge of the backbone to the other.

**Figure 4-7** *Creating Pseudowires*

The following sections describe the different mechanisms used to deliver P2P and MP2MP L2 Ethernet service over MPLS, starting with L2TPv3. You then learn more about Ethernet over MPLS—draft-martini and VPLS.

## PW Setup Via L2TPv3

L2TP provides a dynamic tunneling mechanism for multiple L2 circuits across a packet-oriented data network. L2TP was originally defined as a standard method for tunneling the Point-to-Point Protocol (PPP) and has evolved as a mechanism to tunnel a number of other L2 protocols, including Ethernet. L2TP as defined in RFC 2661, *Layer 2 Tunneling Protocol (L2TP)*, is referred to as L2TPv2. L2TPv3 is an extension of that protocol that allows more flexibility in carrying L2 protocols other than PPP. Notable differences between L2TPv2 and L2TPv3 are the separation of all PPP-related attributes and references and the transition from a 16-bit Session ID and Tunnel ID to a 32-bit Session ID and Control Connection ID, offering more scalability in deploying L2 tunnels.

With L2TPv3 as the tunneling protocol, Ethernet PWs are actually L2TPv3 sessions. An L2TP control connection has to be set up first between two L2TP control connection endpoints (LCCEs) at each end, and then individual PWs can be established as L2TP sessions.

The provisioning of an Ethernet port or Ethernet VLAN and its association with a PW on the PE triggers the establishment of an L2TP session. The following are the elements needed for the PW establishment:

- **PW type**—The type of PW can be either Ethernet port or Ethernet VLAN. The Ethernet port type allows the connection of two physical Ethernet ports, and the Ethernet VLAN indicates that an Ethernet VLAN is connected to another Ethernet VLAN.
- **PW ID**—Each PW is associated with a PW ID that identifies the actual PW.

The entire Ethernet frame without the preamble or FCS is encapsulated in L2TPv3 and is sent as a single packet by the ingress side of the L2TPv3 tunnel. This is done regardless of whether an 802.1Q tag is present in the Ethernet frame. For a PW of type Ethernet port, the ingress side

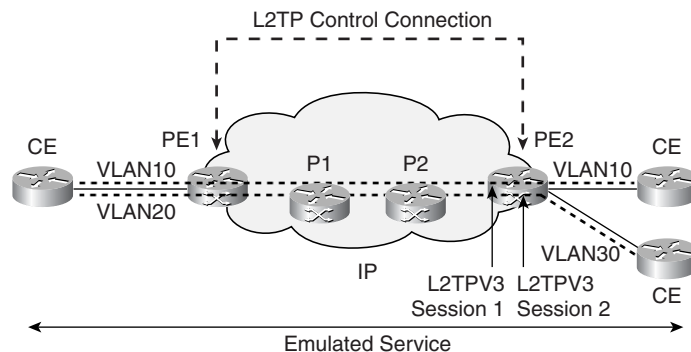
of the tunnel simply de-encapsulates the Ethernet frame and sends it out on the appropriate interface without modifying the Ethernet header. The Ethernet PW over L2TP is homogeneous with respect to packet encapsulation, meaning that both ends of the PW are either VLAN tagged or untagged; however, once the packet leaves the PW, a Native Service Processing (NSP) function within the PE can still manipulate the tag information. For VLAN-to-VLAN connectivity, for example, the egress NSP function may rewrite the VLAN tag if a tag replacement or swapping function is needed.

**NOTE**

The preamble is a pattern of 0s and 1s that tells a station that an Ethernet frame is coming. FCS is the frame check sequence that checks for damage that might have occurred to the frame in transit. These fields are not carried inside the PW.

Figure 4-8 shows an L2TP control connection formed between PE1 and PE2. Over that connection two PWs or L2TPV3 sessions are formed. The two sessions are of type Ethernet VLAN, which means that the PW represents a connection between two VLANs. For session 1, VLAN 10 has been left intact on both sides. For session 2, the NSP function within PE2 rewrites VLAN ID 20 to VLAN ID 30 before delivering the packet on the local segment.

**Figure 4-8** Ethernet over L2TPV3



## Ethernet over MPLS—Draft-Martini

You have seen in the previous section how an Ethernet packet can be transported using an L2TPv3 tunnel over an IP network. The IETF has also defined a way to carry L2 traffic over an MPLS network. This includes carrying Ethernet over MPLS (EoMPLS), Frame Relay, and ATM. This is also referred to as “draft-martini” encapsulation in reference to the author of the original Internet draft that defined Layer 2 encapsulation over MPLS. With this type of encapsulation, PWs are constructed by building a pair of unidirectional MPLS virtual connection (VC) LSPs between the two PE endpoints. One VC-LSP is for outgoing traffic, and

the other is for incoming traffic. The VC-LSPs are identified using MPLS labels that are statically assigned or assigned using the Label Distribution Protocol (LDP).

EoMPLS uses “targeted” LDP, which allows the LDP session to be established between the ingress and egress PEs irrespective of whether the PEs are adjacent (directly connected) or nonadjacent (not directly connected). The following section explains the mechanism of encapsulating the Ethernet frames over the MPLS network and shows two scenarios of using LDP to establish PWs between directly connected and non-directly connected PEs.

## Ethernet Encapsulation

Ethernet encapsulation is very similar to what was described in the “PW Setup Via L2TPv3” section, but a different terminology is introduced. The entire Ethernet frame without any preamble or FCS is transported as a single packet over the PW. The PW could be configured as one of the following:

- **Raw mode**—In raw mode, the assumption is that the PW represents a virtual connection between two Ethernet ports. What goes in one side goes out the other side. The traffic could be tagged or untagged and comes out on the egress untouched.
- **Tagged mode**—In tagged mode, the assumption is that the PW represents a connection between two VLANs. Each VLAN is represented by a different PW and is switched differently in the network. The tag value that comes in on ingress might be overwritten on the egress side of the PW.

The raw and tagged modes are represented in Figure 4-9.

**Figure 4-9** Martini Tunnel Modes

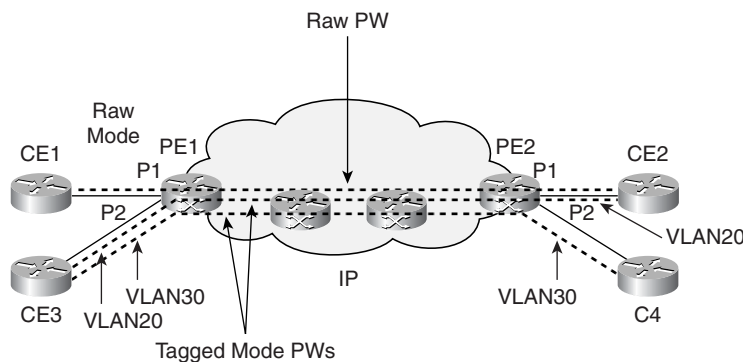


Figure 4-9 shows that PE1 has established a PW of type raw with PE2 over which all traffic coming in on port 1 is mapped. As such, the traffic comes out as-is at the other end of the PW on PE2 port 1. Also, PE1 has defined on port 2 two PWs of type tagged. The first PW maps VLAN 20 on PE1 port 2 and connects it to VLAN 20 on PE2 port 1, and the second PW maps VLAN 30 on PE1 port 2 and maps it to VLAN 30 on PE2 port 2.

JUNIPER Exhibit 1003

App. 6, pg. 105

## Maximum Transmit Unit

Both ends of the PW must agree on their maximum transmission unit size to be transported over the PSN, and the network must be configured to transport the largest encapsulation frames. If MPLS is used as the tunneling protocol, the addition of the MPLS shim layer increases the frame size. If the vendor implementation does not support fragmentation when tunneling the Ethernet service over MPLS, care must be taken to ensure that the IP/MPLS routers in the network are adjusted to the largest maximum transmission unit.

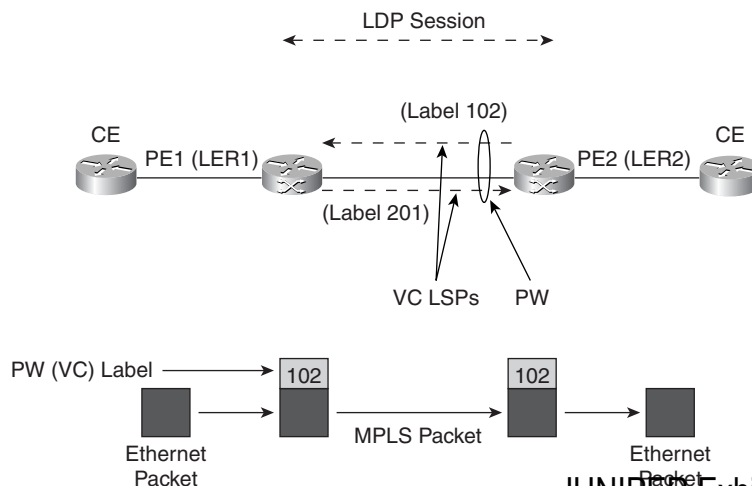
## Frame Reordering

The IEEE 802.3 requires that frames from the same conversation be delivered in sequence. Because the frames are now encapsulated inside PWs, the PW must be able to support frame reordering.

## Using LDP with Directly Connected PEs

Figure 4-10 shows how martini tunnels can be established using LDP between two directly connected PEs, such as PE1 and PE2. First, an LDP session needs to be established between the two PEs. Once the LDP session has been established, all PWs are signaled over that session. In this example, you can see the establishment of one bidirectional PW via two unidirectional VC-LSPs. Once both VC-LSPs are established, the PW is considered operational. PE2 assigns label 102 and sends it to PE1 to be used for propagating traffic from PE1 to PE2. In turn, PE1 assigns label 201 and sends it to PE2 to be used for propagating traffic from PE2 to PE1. The label is pushed into the data packet before transmission, and it indicates to the opposite endpoint what to expect regarding the encapsulated traffic. Remember that this type of encapsulation is used to tunnel not only Ethernet but other types of traffic such as ATM, Frame Relay, and Circuit Emulation traffic. The VC label gives the opposite side an indication of how to process the data traffic that is coming over the VC-LSP.

**Figure 4-10** LDP Between Directly Connected PEs



The VC information is carried in a label mapping message sent in downstream unsolicited mode with a new type of forwarding equivalency class element defined as follows (refer to Figure 4-11):

- **VC Type**—A value that represents whether the VC is of type Frame Relay data-link connection identifier (DLCI), PPP, Ethernet tagged or untagged frames, ATM cell, Circuit Emulation, and so on.
- **PW ID or VC ID**—A connection ID that together with the PW type identifies a particular PW (VC). For P2P tunnels, the VC ID gives an indication of a particular service. You will see in the next section that in the context of an MP2MP VPLS service, the VC ID is indicative of a LAN.
- **Group ID**—Represents a group of PWs. The Group ID is intended to be used as a port index or a virtual tunnel index. The Group ID can simplify configuration by creating a group membership for all PWs that belong to the same group, such as an Ethernet port carrying multiple PWs.
- **Interface Parameters**—A field that is used to provide interface-specific parameters, such as the interface maximum transmission unit.

**Figure 4-11** LDP Forwarding Equivalency Class

//	PW Type (VC Type)	//
Group ID		
PW ID (VC ID)		
Interface Parameters		

MPLS PWs are formed using two unidirectional VC-LSPs, which means that for each PW that is established from ingress to egress, a “matching” PW needs to be established between egress and ingress with the same PW ID and PW type.

In the remainder of this book, the terms PW and VC-LSP are used interchangeably, but remember that a PW is formed of two unidirectional VC-LSPs, one inbound and one outbound.

## Non-Directly Connected PEs

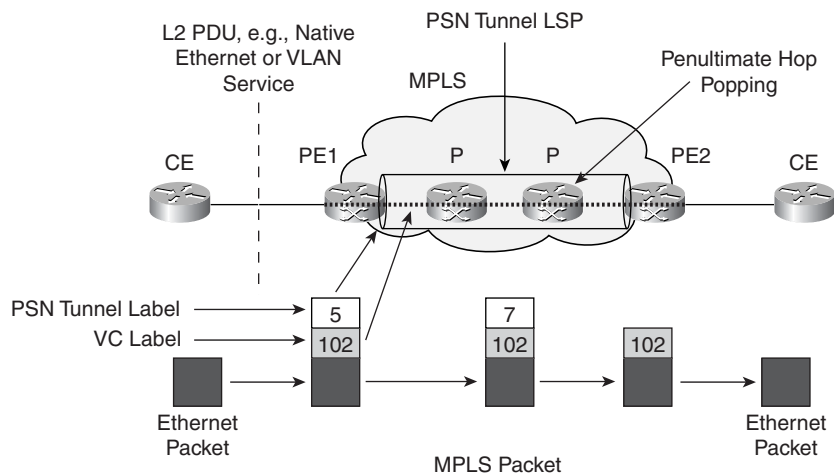
If the PEs are not directly connected, the PE-to-PE traffic has to traverse the MPLS backbone across P core routers. These routers do not need to get involved with the different services offered at the edge and are concerned only with transporting the traffic from PE to PE. To hide the information from the P routers, LSP tunnels are constructed between the different PEs using targeted LDP, and the different PWs can share these tunnels. The construction of the LSP

tunnels does not relate to the Ethernet MPLS service whatsoever. These tunnels can be constructed via different methods, such as GRE, L2TP, or MPLS. If constructed via MPLS, a signaling protocol such as RSVP-TE can be constructed to traffic-engineer these LSP tunnels across the network (RSVP-TE is explained in Chapter 6).

In Figure 4-12, an LSP tunnel, called a packet-switched network (PSN) tunnel LSP, is constructed between PE1 and PE2, and the PW is carried across that tunnel. The PSN tunnel LSP is constructed by having PE1 push a tunnel label that gets the packets from PE1 to PE2. The PSN tunnel label is pushed on top of the VC label, which gives the other side an indication of how to process the traffic. The P routers do not see the VC label and are only concerned with switching the traffic between the PE routers irrespective of the service (indicated by the VC labels) that is carried. The following describes the process of transporting a packet from ingress PE to egress PE:

- 1 When PE1 sends a Layer 2 PDU to PE2, it first pushes a VC label on its label stack and then a PSN tunnel label.
- 2 As shown in Figure 4-12, a targeted LDP session is formed between PE1 and PE2.

**Figure 4-12** LDP Between Non-Directly Connected PEs



- 3 PE2 gives PE1 label 102 to be used for traffic going from PE1 to PE2 (the same scenario happens in the reverse direction).
- 4 Label 102 is pushed by PE1, and then a PSN tunnel LSP label 5 is pushed on top.
- 5 The P routers use the upper label to switch the traffic toward PE2. The P routers do not have visibility to the VC labels.

- 6 The last router before PE2 performs a penultimate hop popping function to remove the upper label before it reaches PE2. *Penultimate hop popping* is a standard MPLS function that alleviates the router at the end of the LSP (PE2 in this case) from performing a popping function and examining the traffic beneath at the same time. PE2 receives the traffic with the inner label 102, which gives an indication of what is expected in the PW.

So far you have seen a P2P L2 service over MPLS. Next, MP2MP is discussed when a LAN is emulated over MPLS using VPLS.

## Virtual Private LAN Service

With Virtual Private LAN Service, an L2VPN emulates a LAN that provides full learning and switching capabilities. Learning and switching are done by allowing PE routers to forward Ethernet frames based on learning the MAC addresses of end stations that belong to the VPLS. VPLS allows an enterprise customer to be in full control of its WAN routing policies by running the routing service transparently over a private or public IP/MPLS backbone. VPLS services are transparent to higher-layer protocols and use L2 emulated LANs to transport any type of traffic, such as IPv4, IPv6, MPLS, IPX, and so on.

VPLS is flexible because it emulates a LAN, but by doing so it has all the limitations of Ethernet protocols, including MAC addresses, learning, broadcasts, flooding, and so on. The difference between VPLS and EoMPLS is that VPLS offers an MP2MP model instead of the previously discussed P2P model with L2TPv3 or EoMPLS using martini tunnels.

With VPLS, the CEs are connected to PEs that are VPLS-capable. The PEs can participate in one or many VPLS domains. To the CEs, the VPLS domains look like an Ethernet switch, and the CEs can exchange information with each other as if they were connected via a LAN. This also facilitates the IP numbering of the WAN links, because the VPLS could be formed with a single IP subnet. Separate L2 broadcast domains are maintained on a per-VPLS basis by PEs. Such domains are then mapped into tunnels in the service provider network. These tunnels can either be specific to a VPLS (for example, IP tunnels) or shared among several VPLSs (for example, with MPLS LSPs).

The PE-to-PE links carry tunneled Ethernet frames using different technologies such as GRE, IPsec, L2TP, MPLS, and so on. Figure 4-13 shows a typical VPLS reference model.

As Figure 4-13 shows, MPLS LSP tunnels are created between different PEs. These MPLS tunnels can be shared among different VPLS domains and with other services such as EoMPLS tunnels, Layer 3 MPLS VPN tunnels, and so on. The PE routers are configured to be part of one, many, or no VPLS, depending on whether they are participating in a VPLS service.

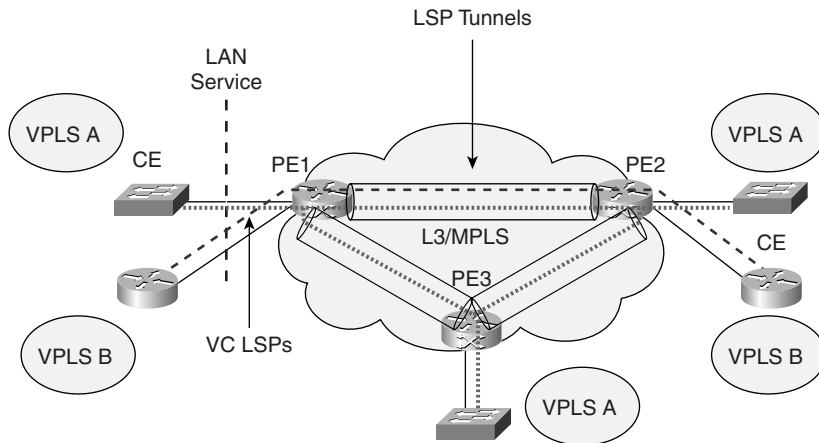
---

**NOTE**

The access network connecting the CEs to the PEs could be built with Ethernet technology or with next-generation SONET/SDH running Ethernet framing over the Generic Framing Protocol (GFP) or any logical links such as ATM PVCs or T1/E1 TDM or any virtual or physical connections over which bridged Ethernet traffic is carried.



Figure 4-13 VPLS Reference Model



The following sections discuss the different aspects of a VPLS model:

- VPLS requirements
- Signaling the VPLS service
- VPLS encapsulation
- Creating a loop-free topology
- MAC address learning
- MAC address withdrawal
- Unqualified versus qualified learning
- Scaling the VPLS service via hierarchical VPLS (HVPLS)
- Autodiscovery
- Signaling using BGP versus LDP
- Comparing the Frame Relay and MPLS/BGP approaches
- L2VPN BGP model
- Frame Relay access with MPLS edge/core
- Decoupled Transparent LAN Service (DTLS)

## VPLS Requirements

Following are the basic requirements of a VPLS service:

- **Separation between VPLS domains**—A VPLS system must distinguish different customer domains. Each customer domain emulates its own LAN. VPLS PEs must maintain a separate virtual switching instance per VPN.

JUNIPER Exhibit 1003

App. 6, pg. 110

- **MAC learning**—A VPLS should be capable of learning and forwarding based on MAC addresses. The VPLS looks exactly like a LAN switch to the CEs.
- **Switching**—A VPLS switch should be able to switch packets between different tunnels based on MAC addresses. The VPLS switch should also be able to work on 802.1p/q tagged and untagged Ethernet packets and should support per-VLAN functionality.
- **Flooding**—A VPLS should be able to support the flooding of packets with unknown MAC addresses as well as broadcast and multicast packets. Remember that with Ethernet, if a switch does not recognize a destination MAC address, it should flood the traffic to all ports within a certain VLAN. With the VPLS model shown in Figure 4-13, if a VPLS-capable device receives a packet from VPLS A with an unknown MAC destination address, the VPLS device should replicate the packet to all other VPLS-capable devices that participate in VPLS A.
- **Redundancy and failure recovery**—The VPLS should be able to recover from network failure to ensure high availability. The service should be restored around an alternative path, and the restoration time should be less than the time the CEs or customer L2 control protocols need to detect the failure of the VPLS. The failure recovery and redundancy of MPLS depends on how fast MPLS paths can be restored in case of a failure and how fast the network can stabilize. Chapter 6 discusses MPLS fast restoration.
- **Provider edge signaling**—In addition to manual configuration methods, VPLS should provide a way to signal between PEs to auto-configure and to inform the PEs of membership, tunneling, and other relevant parameters. Many vendors have adopted LDP as a signaling mechanism; however, there are some who prefer BGP as used in RFC 2547, *BGP/MPLS VPNs*.
- **VPLS membership discovery**—The VPLS control plane and management plane should provide methods to discover the PEs that connect CEs forming a VPLS. Different mechanisms can be used to achieve discovery. One method is via the use of BGP, as adopted in the L3VPN model. However, there is some disagreement in the industry on whether BGP implementations are appropriate, due to the complexity of BGP and the fact that it cannot signal a different label to each VPLS peer, as required by MAC learning. A proposal for using BGP promotes the use of block label distribution, as explained in the “DTLS—Decoupling L2PE and PE Functionality” section later in this chapter.
- **Interprovider connectivity**—The VPLS domain should be able to cross multiple providers, and the VPLS identification should be globally unique.
- **VPLS management and operations**—VPLS configuration, management, and monitoring are very important to the success of the VPLS service. Customer SLAs should be able to be monitored for availability, bandwidth usage, packet counts, restoration times, and so on. The metrics that have been defined by the MEF regarding performance and bandwidth parameters should apply to the VPLS service.

## Signaling the VPLS Service

Signaling with VPLS is the same as described in the section “Ethernet over MPLS—Draft-Martini,” with LDP using a forwarding equivalency class element. The main difference is that in the P2P martini tunnel, the VC ID is a service identifier representing a particular service on the Ethernet port, such as a different P2P VLAN. With VPLS, the VC ID represents an emulated LAN segment, and its meaning needs to be global within the same provider and across multiple providers.

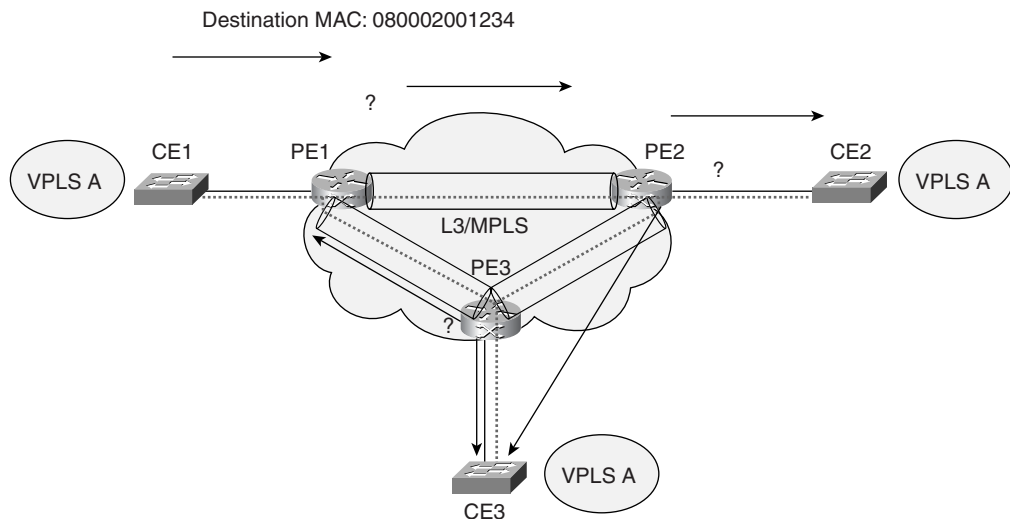
## VPLS Encapsulation

VPLS encapsulation is derived from the martini encapsulation used for a P2P EoMPLS service. The packet is always stripped from any service-related delimiter that is imposed by the local PE. This ensures that the Ethernet packet that traverses a VPLS is always a customer Ethernet packet. Any service delimiters, such as VLAN or MPLS labels, can be assigned locally at the ingress PE and stripped or modified in the egress PE.

## Creating a Loop-Free Topology

The problem with having a VPLS domain emulate a LAN is that it can create the same circumstances that create a loop in a LAN. With L2 Ethernet networks, Spanning Tree Protocol is used to prevent loops caused by the L2 flooding mechanism. In the case of VPLS, the same scenario could happen as illustrated in Figure 4-14.

**Figure 4-14** *L2 Loops*

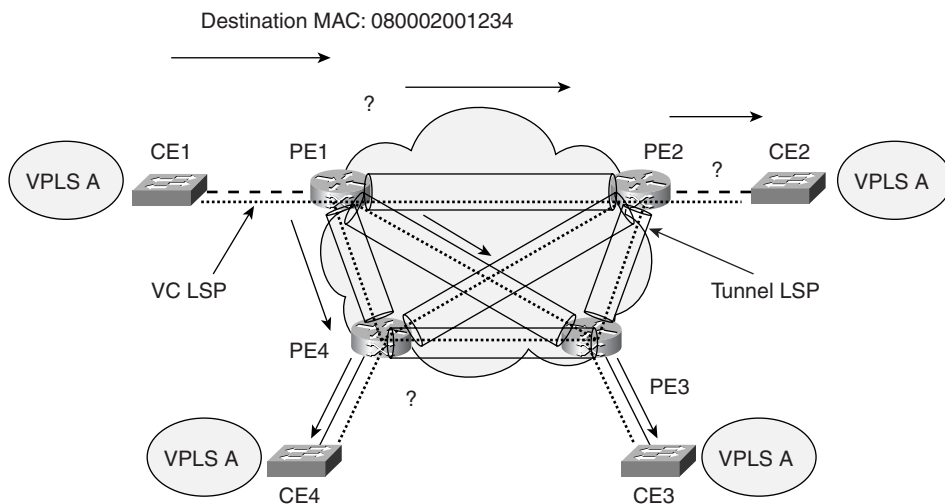


In Figure 4-14, three LSP tunnels connect PE1, PE2, and PE3. VPLS A is emulating a LAN that is carried over these LSP tunnels. If CE1 sends a packet with a destination MAC address, say 080002001234, that is unknown by PE1, PE1 has to flood, or replicate, that packet over the two tunnels connecting it to PE2 and PE3 that participate in the same VPLS. If PE2 does not know of the destination as well, it sends the packet to PE3 and CE2. In the same manner, if PE3 does not know of the destination, it sends the packet to PE1 and CE3, and the loop continues. To break the loop, Spanning Tree Protocol (STP) has to run on the PEs, and in the same way that Ethernet frames are tunneled over the MPLS LSPs, STP BPDUs also have to be tunneled.

To avoid the deployment of spanning trees, a full mesh of LSPs needs to be installed between the PEs, and each PE must support a split-horizon scheme wherein the PE must not forward traffic from one PW to another in the same VPN. This works because each PE has direct connectivity to all other PEs in the same VPN.

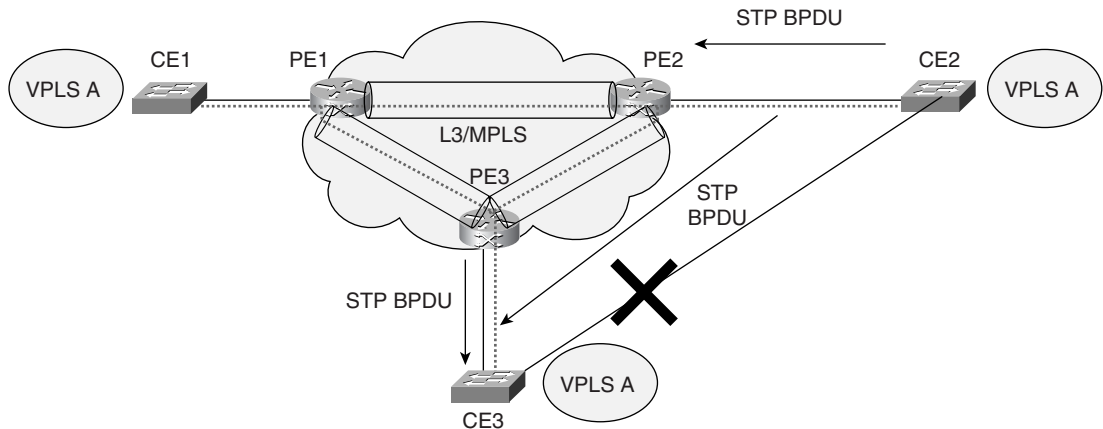
In Figure 4-15, a full mesh of tunnel LSPs and VC-LSPs (which are used to demultiplex the service over the tunnel LSPs) has been configured between all PEs. A PE receiving a packet over a VC-LSP cannot forward that packet to other VC-LSPs. PE1 receives an Ethernet packet with an unknown destination and replicates that packet over the three VC-LSPs that connect it to PE2, PE3, and PE4. Because there is a full mesh, PE2, PE3, and PE4 assume that the same packet they received has already been sent by the other PEs and thus do not replicate it. This prevents loops from taking place. Requiring a full mesh of LSPs becomes an issue if the PE functionality is moved closer into the access cloud, such as in the basement of multitenant unit (MTU) buildings. This would create an explosion of an LSP mesh that does not scale. The section “Scaling the VPLS Service Via Hierarchical VPLS” later in this chapter explains how such a scenario is solved.

**Figure 4-15** *Avoiding Loops Via Full Mesh*



In some cases, an enterprise customer can create a backdoor loop by connecting multiple sites directly via an L2 connection. To avoid loops, STP can be run on the CEs, and the STP BPDUs are tunneled over the MPLS cloud like any other data packets. This is shown in Figure 4-16.

**Figure 4-16** *Backdoor Loops*



In Figure 4-16, CE2 and CE3 have a direct L2 connection, creating a loop, because traffic between CE2 and CE3 can traverse two different paths: the direct connection and the MPLS cloud. If the customer runs STP between the two CEs, STP BPDUs can be tunneled over the MPLS cloud, causing the loop to break. In case the MPLS connection or the direct connection fails, traffic is switched over the remaining connection.

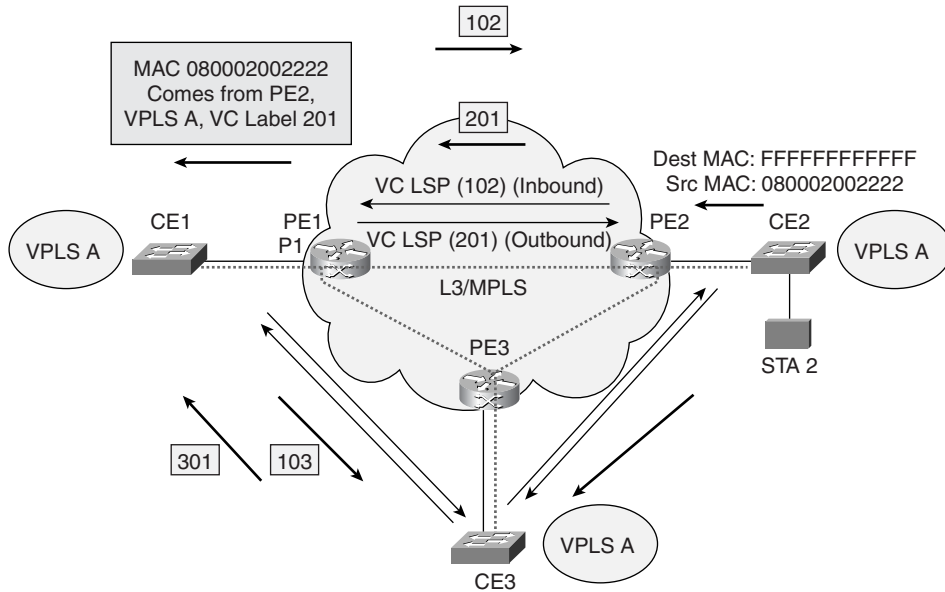
So, there are many ways to create loops in an Ethernet L2 switched architecture. Some of these loops can be caused by the service provider equipment and some by the enterprise equipment itself. This creates more strain on the operation and management of these systems, when problems occur. Well-defined rules need to be set to indicate which L2 control PDUs are to be carried over the provider network. This prevents finger-pointing between the customer and provider in case problems such as broadcast storms occur.

## MAC Address Learning

In the P2P PW scenarios in which Ethernet packets come in on one side of the PW and come out on the other side, MAC learning is not necessary. What goes in the tunnel comes out on the other end. VPLS operates in any-to-any MP2MP mode. This means that a PE is connected with multiple VC-LSPs to different PEs that participate in the multiple VPLS domains, and the PE needs to decide which LSPs to put the traffic on. This decision is based on destination MAC addresses that belong to a certain VPLS. It is unreasonable to assume that this function can be statically configured (although it could), because many MAC addresses would need to be mapped to many LSPs. MAC learning allows the PE to determine from which physical port or LSP a particular MAC address came.

Figure 4-17 shows an example of how MAC learning is achieved.

**Figure 4-17** MAC Learning



In Figure 4-17, PE1, PE2, and PE3 establish pairs of VC-LSPs between each other as follows:

- 1 Using LDP, PE2 signals VC label 201 to PE1, and PE1 signals VC label 102 to PE2.
- 2 A station behind CE2, STA 2, with a MAC address of 080002002222, sends a broadcast packet to destination MAC FFFFFFFF. PE2 recognizes that STA 2 belongs to VPLS A (via configuration or other mechanism) and replicates the packet to the two VC-LSPs connected to PE1 and PE3 that also participate in VPLS A.
- 3 When the packet comes to PE1 on PE1’s inbound VC-LSP, it associates the MAC address of STA 2 with the “outbound” VC-LSP in the same VC-LSP pair that constitutes the PW between PE2 and PE1.
- 4 From then on, if PE1 receives a packet destined for MAC 080002002222, it automatically sends it on its outbound VC-LSP using label 201 (which was signaled by PE2).

This process constitutes MAC learning on the VC-label side. Standard Ethernet MAC learning occurs on the Ethernet port side, where PE2, for example, associates MAC 080002002222 with its local Ethernet port or the VLAN it came on. This process continues until all PEs have learned all MAC addresses on their local ports/VLANs and across the MPLS cloud. Notice that PE1 signals two different labels to PE2 and PE3. In this example, PE1 signals label 102 to PE2 and 103 to PE3. This way, PE1 can distinguish inbound packets from PE2 and PE3.

## MAC Address Withdrawal

L2 Ethernet switching includes a mechanism called MAC aging that lets MAC addresses be aged out of an Ethernet switch MAC table after a certain period of inactivity. In some cases, such as an MTU building that is dual-homed to two different Ethernet switches in the central office (CO), faster convergence can occur if a mechanism exists to age out (withdraw) or relearn MAC addresses in a way that is faster than the traditional L2 MAC aging. The IETF has defined a MAC type length value (TLV) field that can be used to expedite learning of MAC addresses as a result of topology change.

## Unqualified Versus Qualified Learning

When a PE learns MAC addresses from the attached customers, these MAC addresses are kept in a Forwarding Information Base (FIB). The FIB should keep track of the MAC addresses and on which PWs they were learned. This allows MAC addresses to be tracked by VPLS. This is different from the traditional MAC learning of Ethernet switches, where all MAC addresses are shared by a single customer. VPLS can operate in two learning modes, unqualified and qualified.

In unqualified learning, a customer VPLS is a port-based service where the VPLS is considered a single broadcast domain that contains all the VLANs that belong to the same customer. In this case, a single customer is handled with a single VPLS. On the other hand, qualified learning assumes a VLAN-based VPLS where each customer VLAN can be treated as a separate VPLS and as a separate broadcast domain. The advantage of qualified learning is that customer broadcast is confined to a particular VLAN.

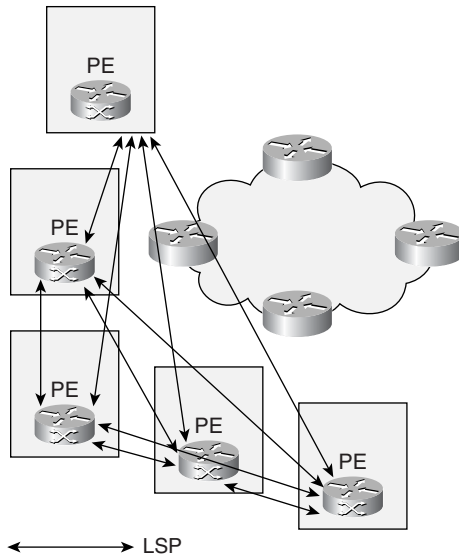
## Scaling the VPLS Service Via Hierarchical VPLS

The VPLS service requires a full mesh of VC-LSPs between the PE routers. This works adequately if the PE routers are contained in COs and the different customers are aggregated in these COs. In the case of MTU deployments, the PEs are deployed in the building basements where multiple customers are aggregated. In this case, starting the VPLS service in the PE might cause scalability problems because there are many more buildings than COs. A full mesh of LSPs between all the buildings that participate in the VPLS service would cause an unmanageable LSP deployment. For  $x$  PEs that are deployed,  $x * (x - 1) / 2$  bidirectional LSPs need to be deployed. Remember also that it takes two LSPs—one inbound and one outbound—to construct a bidirectional PW, which means that  $x * (x - 1)$  unidirectional VC-LSPs need to be signaled.

Figure 4-18 shows a deployment in which the VPLS starts in the basement of MTU buildings and a full mesh of LSPs is required between PEs. This LSP explosion will cause an operational nightmare.

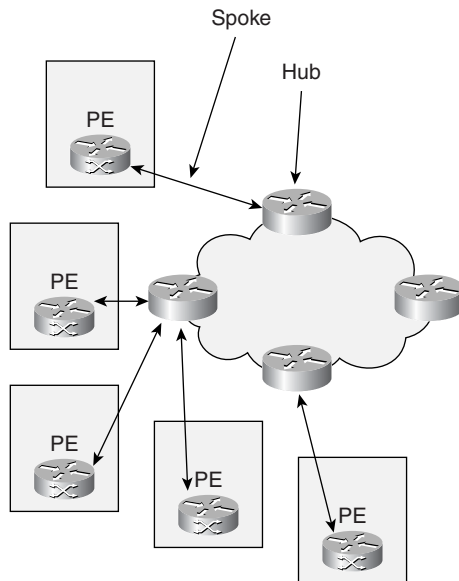
For any “new” MTU building that is added to the network, the new MTU must be meshed to every PE in the existing MTUs, which doesn’t scale. Packets get flooded over all LSPs participating in a VPLS; if the MAC destination is unknown, this puts a big load on the MTU PE.

Figure 4-18 Full Mesh



A better approach for MTUs is to create a hierarchical VPLS (HVPLS) model in which the MTU PEs establish access tunnels (spokes) to the CO PEs, and the CO PEs (hubs) establish a full mesh. This is shown in Figure 4-19.

Figure 4-19 HVPLS





The hierarchical model scales better, because a new MTU that is added to the network has to establish an LSP only with the local PE and does not need to establish LSPs with every other PE. This is a major operational cost saving.

There are many flavors for the MTU and the CO PEs. The IETF has adopted the following terminology that is used in the rest of the chapter:

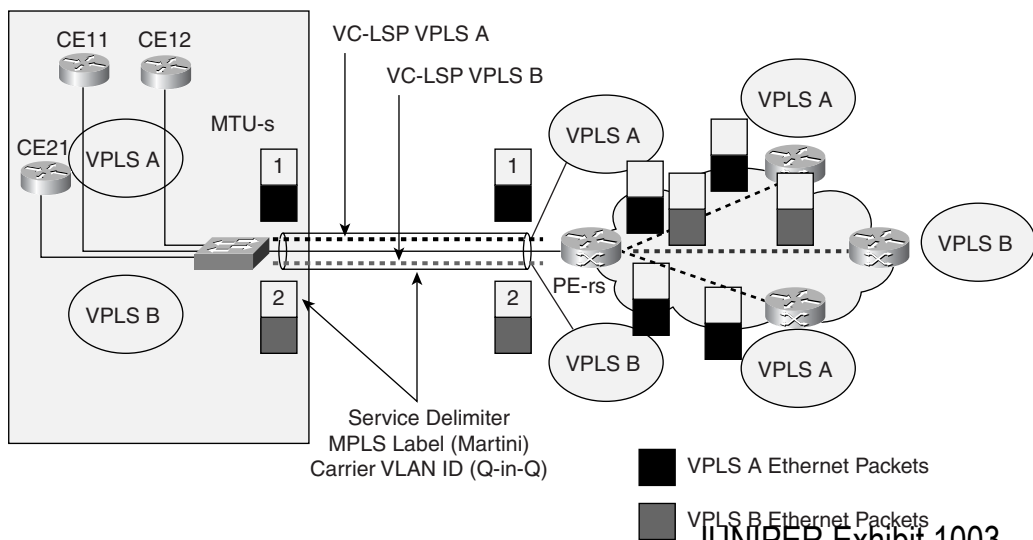
- **MTU-s**—This is a PE that is placed in the MTU and is capable of doing MAC learning and L2 switching/bridging. This could be a pure L2 Ethernet switch, or an L2 Ethernet switch that is capable of MPLS tagging and forwarding but does not have to do any IP routing.
- **PE-r**—This is a PE that is capable of IP routing/MPLS but is not capable of MAC learning. This device can be placed in the MTU or the CO. This is basically an IP/MPLS router.
- **PE-rs**—This a PE that is capable of both L2 switching and IP routing.

The following sections explain two different scenarios used in service provider deployments: one for MTU-s deployments and the other for PE-rs.

### MTU Device Supports MAC Learning and L2 Switching (MTU-s)

In this scenario, the MTU-s is an L2 Ethernet switch that is capable of MAC learning and can do switching based on MAC addresses. The MTU-s does all the normal bridging functions of learning and replications on all its ports, including the virtual spoke ports, which are the PWs that connect the MTU-s to the PE-rs. The ability of the MTU-s to do MAC learning and bridging simplifies the signaling between the MTU-s and the PE-rs at the CO, because the MTU-s can associate all the access ports belonging to the same VPLS with a single PW between the MTU-s and the PE-rs. This is better illustrated in Figure 4-20.

**Figure 4-20** Sample MTU-s



In Figure 4-20, a service provider is offering a VPLS service to two customers via an MTU-s in the basement of the building. The MTU-s connects to a PE-rs in the CO. Note that on the MTU-s, two access Ethernet ports are assigned to VPLS A. These access ports are connected to CE11 and CE12, which both connect to the same customer. This scenario could occur if the same customer has two different locations that are in the same MTU or a nearby building and all connections are serviced by the same MTU-s toward the CO. In this case, the MTU-s switches all traffic between CE11 and CE12 locally via regular L2 switching and switches traffic between CE11 and CE12 and the remote sites in VPLS A using a single PW between the MTU-s and the PE-rs in the CO.

Because the MTU-s also services VPLS B, the service provider has to assign a service delimiter to traffic coming from VPLS A and VPLS B to differentiate between the two customers. Remember that this was discussed in Chapter 3, “Metro Ethernet Services,” which introduced the concept of using carrier VLAN IDs to differentiate between customer traffic. In this example, you could set up the spoke in two ways:

- The service provider is using Q-in-Q to separate the customer traffic on the MTU-s and to indicate to the PE-rs which traffic belongs to which VPLS. In this case, the service delimiter is a carrier VLAN ID carried on top of the customer’s Ethernet packet. The customer traffic itself also could carry customer-specific VLAN tags; however, those tags are not seen by the service provider.
- The service provider is using two martini EoMPLS PWs to carry traffic from the different customers. In this case, the MPLS tag on top of the customer’s Ethernet traffic is the service delimiter recognized by the PE-rs.

The decision of whether to use Q-in-Q or martini tunnels depends on the equipment the vendor uses in the MTU and the CO. In some cases, the MTU equipment doesn’t support MPLS. In other cases, the MTU and CO equipment does not interoperate when using Q-in-Q. You should also remember that some Ethernet switch vendors support neither VLAN stacking on a per-customer basis nor MPLS. You should not use such equipment in MTU deployments.

Notice in this example that the PWs used between the MTU-s and the PE-rs have achieved multiple functions:

- The need for full PW mesh between the MTU-rs is eliminated. Only one PW is used per VPLS.
- The signaling overhead is minimized because fewer PWs are used.
- MTU-s devices are only aware of the PE-rs they attach to and not to all MTU-s devices that participate in the VPLS.
- An addition of a new MTU-s does not affect the rest of the network.

The MTU-s learns MAC addresses both from the Ethernet customer connections in the building and from the spoke PWs. The MTU-s associates the MAC addresses per VPLS. If an MTU-s receives a broadcast packet or a packet with an unknown destination MAC, the packet is flooded (replicated) over all the MTU-s physical or logical connections that participate within

the VPLS. Note that there is one PW per VPLS on the spoke connection, so the packet is replicated only once per VPLS.

The MTU-s device and the PE-rs device treat each spoke connection like a physical port on the VPLS service. On the physical ports, the combination of the physical port and VLAN tag is used to associate the traffic with a VPLS instance. On the spoke port, the VC label or carrier VLAN ID (for Q-in-Q) is used to associate the traffic with a particular VPLS. L2 MAC address lookup is then used to find out which physical port the traffic needs to be sent on.

The PE-rs forms a full mesh of tunnels and PWs with all other PE-rs devices that are participating in the VPLS. A broadcast/multicast or a packet with an unknown MAC destination is replicated on all PWs connected to the PE-rs for a certain VPLS. Note that the PE-rs can contain more VPLS instances than the MTU-s, because the PE-rs participates in all the VPLSs of the MTU buildings that are attached to it, while the MTU-s only participates in the VPLS of the customers in a particular building. Also, the MAC learning function is done twice: once at the MTU-s and another time at the PE-rs.

### PE-rs Issues with MAC Learning

The fact that the PE-rs is doing MAC learning raises concerns with service providers. The PE-rs has to learn all the MAC addresses that exist in all VPLS instances it participates in. This could be in the hundreds of thousands of MAC addresses that need to be learned if the VPLS service is delivering LAN connectivity between CEs that are L2 switches. Remember that a VPLS emulates a LAN service and learns all MAC addresses it hears from all stations connected to the LAN. If the CEs are L2 Ethernet switches, the VPLS will learn all MAC addresses behind the Ethernet switch. Some of these concerns can be alleviated through different approaches:

- If the CE equipment is an IP router, the VPLS learns only the MAC addresses of the IP router interfaces that are connected to the VPLS. MAC stations behind IP routers are hidden, because IP routers route based on IP addresses and not MAC addresses. In this model, the MAC address space is very manageable.
- If the CEs are L2 switches, it is possible to use filtering mechanisms on the MTU-s to allow service for only a block of the customer's MAC addresses and not all of them. Filtering helps reduce the explosion of MAC addresses on the PE-rs; however, it adds more management overhead for both the customer and the service provider.

A different model can be used to allow the MTU-s to do MAC learning at the building and not to do MAC learning at the PE-rs. This model is called the *Decoupled Transparent LAN Service (DTLS)*, which is explained later in this chapter in the section “DTLS—Decoupling L2PE and PE Functionality.”

### Non-Bridging Devices as Spokes

In some cases, existing IP routers are deployed as spokes. As previously described, the IETF calls such a device a PE-r, to indicate routing functionality only. These routers are not capable of bridging and cannot switch packets based on MAC addresses. To offer an L2 service using

the PE-r, it is possible to create PWs between the PE-r and the CO PE-rs, where all the L2 switching functions are done at the CO. This model creates more overhead, because unlike the MTU-s, where all access ports belonging to the same VPLS are mapped to a single PW, the PE-r requires that each access port is mapped to its own PW. This is illustrated in Figure 4-21.

**Figure 4-21** *Spoke Device Is a Router*

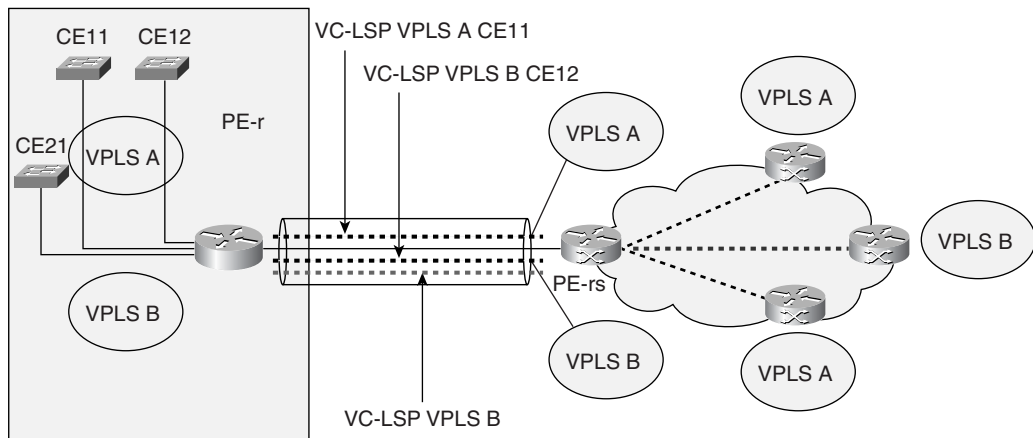


Figure 4-21 uses a PE-r as a spoke. Note that VPLS A now requires two PWs—one for CE11 and one for CE12—that belong to the same customer. For any traffic that needs to be switched between the two access ports of the same customer that are connected to CE11 and CE12, that traffic needs to be transported to the CO and switched at the PE-rs.

### Dual-Homed MTU Devices

It is possible to dual-home an MTU device to protect against the failure of a spoke or the failure of a PE-rs at the CO. *Dual-home* refers to connecting the MTU device via two separate spokes.

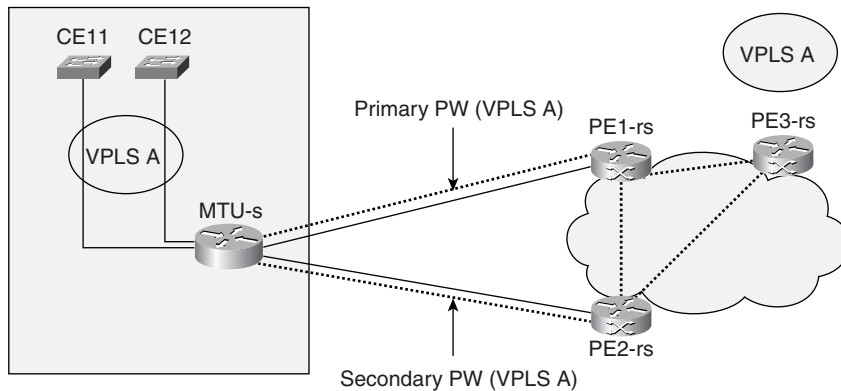
Figure 4-22 shows an MTU-s device that is dual-homed to the PE-rs at the CO via two PWs, one primary and one backup. To prevent an L2 loop in the network, the primary PW is active and passing traffic while the secondary PW is inactive. In this scenario, spanning tree is not needed, because only a single PW is active at the same time. In normal operation, all PE-rs devices participating in VPLS A learn the MAC addresses behind MTU-s via the primary PW connected to PE1-rs. The following two scenarios might take place:

- Failure of the primary PW**—In this case the MTU-s immediately switches to the secondary PW. At this point the PE2-rs that is terminating the secondary PW starts learning MAC addresses on the spoke PW. The speed of convergence in the network depends on whether MAC TLVs are used, as described in the “MAC Address Withdrawal” section earlier in this chapter. If the MAC address TLVs are used, PE2-rs sends a flush message to all other PE-rs devices participating in the VPLS service. As such,

all PE-rs devices converge on PE2-rs to learn the MAC addresses. If the MAC TLV is not used, the network is still operational and converges using the traditional L2 MAC learning and aging. During this slow convergence, the PE-rs devices slowly learn the MAC addresses in the network.

- **Failure of the PE1-rs**—In this case, all PWs that are terminated at PE1-rs fail, and the network converges toward PE2-rs.

**Figure 4-22** *Dual-Homed MTU Device*



## Autodiscovery

*Autodiscovery* refers to the process of finding all the PEs that participate in a given VPLS. So far, this chapter has assumed that this function is manual, meaning that the network operator dedicates certain PEs to belong to a certain VPLS and configures that information on each PE belonging to the VPLS. This process can be configuration-intensive, especially with a large number of PEs, because manual configuration and deletion are needed every time a PE is added to or removed from the network. With autodiscovery, each PE discovers which other PEs are part of the same VPLS and discovers PEs when they are added to or removed from the network. Different mechanisms have been proposed by different vendors, such as the use of BGP, LDP, or DNS to achieve autodiscovery. This section elaborates on BGP and how it compares with LDP.

BGP uses the concept of extended communities to identify a VPLS. PEs exchange information via direct Internal BGP (IBGP) or External BGP (EBGP) peering or route reflectors.

You saw at the beginning of this chapter that BGP is used with MPLS L3VPNs to achieve discovery of VPN information. A similar approach is used to achieve VPLS autodiscovery by having the routes exchanged in BGP carry a VPN-L2 address. A VPN-L2 address contains a route distinguisher (RD) field that distinguishes between different VPN-L2 addresses. Also, a BGP route target (RT) extended community is used to constrain route distribution between PEs. The RT is indicative of a particular VPLS. Because a PE is fully meshed with all other PEs, it receives BGP information from all PEs. The PE filters out the information based on the route target and learns only information pertinent to the route targets (VPLSs) it belongs to.

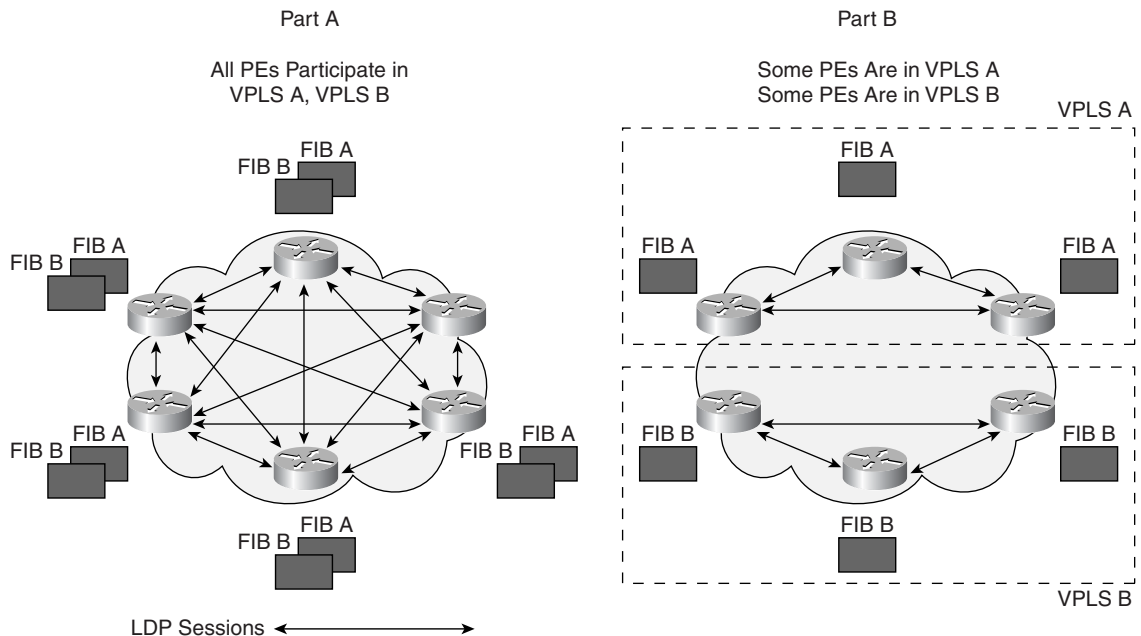
## Signaling Using BGP Versus LDP

In this chapter, you have learned about the use of LDP as a signaling mechanism to establish and tear down PWs between PEs. Some vendors have adopted BGP as a signaling mechanism because of its scalability and its ability to support VPLS deployment across multiple providers. This section presents a more detailed comparison of the use of LDP and BGP as a signaling mechanism and BGP.

With LDP used as the signaling protocol, targeted LDP sessions are established between PE peers. An LDP session is called “targeted” because it is set directly between two PEs that do not have to be adjacent. These PEs exchange MPLS labels directly, and that information is hidden from the routers that exist on the path between these PE peers. You have seen that a full mesh of these peers between PEs is needed per VPLS. If all PE routers participate in every VPLS, a full mesh is needed between all PEs. Also, each PE needs to carry a separate FIB per VPLS, which increases the number of FIBs per PE. However, it is possible to segment the network into PEs that have separate VPLS coverage, meaning that they do not serve a common set of VPLSs. In this case, the LDP mesh is needed only between the PEs covering a particular VPLS, and the signaling and the number of FIBs per PE are reduced.

If all PEs participate in all VPLS instances, there is a full LDP mesh between all PEs, and each PE carries a FIB per VPLS, as shown in Figure 4-23, Part A. Figure 4-23, Part B, shows that three PEs participate in VPLS A and carry a VPLS A FIB (FIB A) while the other PEs carry a VPLS B FIB. Note that a full mesh between all PEs is not required.

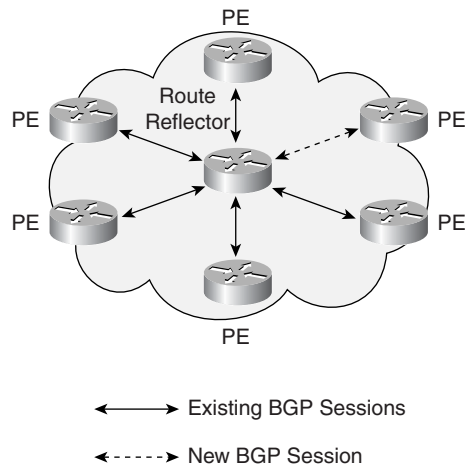
**Figure 4-23** LDP Signaling Options



Vendors proposing BGP as a signaling mechanism between PEs argue that BGP offers more scalability and is already proven to work for L3VPNs as defined in RFC 2547. Also, BGP can be used for both signaling and PE discovery, whereas LDP is used only for signaling. BGP uses what is called a *route reflector* to solve the full-mesh PE-to-PE session issue and the fact that, with LDP, every time a new PE is added to the network, a full mesh needs to be established with all PEs (in the same VPLS). The route reflector concept allows PEs to operate in a client/server model, where the PEs peer with a single or multiple route reflectors (for redundancy), and the route reflector relays information between the different PEs. In this case, if a new PE is added to the network, that PE needs to establish only a single peering session with the route reflector.

Figure 4-24 shows all PEs peering with a route reflector. A new PE added to the network has to peer only with the route reflector.

**Figure 4-24** Signaling Via BGP with Route Reflectors



On the other hand, using BGP does create the issue of requiring label ranges, because BGP cannot direct label mappings to a specific peer. The use of label ranges is covered in the upcoming section “L2VPN BGP Model.”

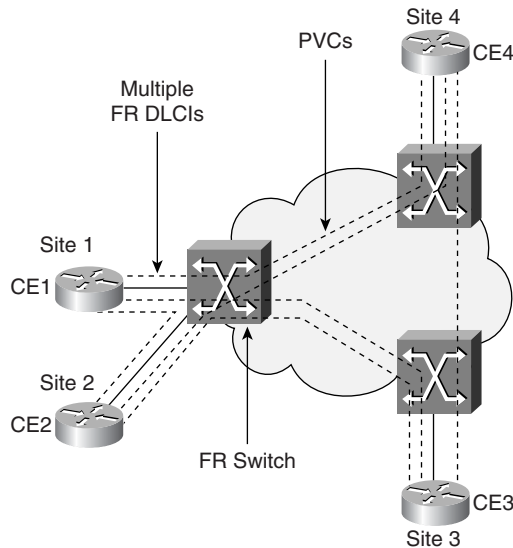
### Comparison Between the Frame Relay and MPLS/BGP Approaches

This section first briefly compares Frame Relay VPNs and MPLS L2VPNs and then delves into a discussion about how some IETF drafts proposing BGP are influenced by the Frame Relay VPN model.

As Figure 4-25 shows, a Frame Relay VPN with any-to-any connectivity between the different sites requires a full mesh of PVCs between the different CEs. The network uses Frame Relay

at the access and Frame Relay/ATM at the edge and core. The physical connection between the CE and the Frame Relay network is assigned multiple DLCIs, and each DLCI is used to switch the traffic from one CE all the way to another CE.

**Figure 4-25** *Frame Relay Access, Edge, and Core*



Such networks have two drawbacks. First, the whole network is locked into a single technology, such as Frame Relay or ATM. Second, adding a new site into the VPN and connecting that site to the rest of the VPN causes an operational headache because many PVCs need to be configured site to site.

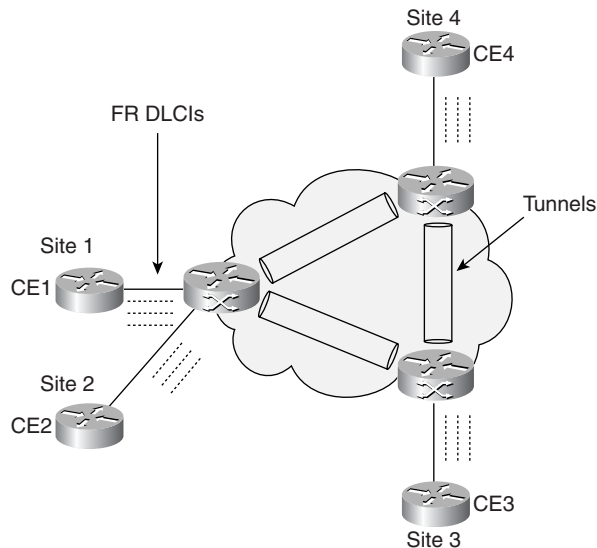
Figure 4-26 shows the same VPN but with an MPLS deployment at the core and with the possibility of using Frame Relay, Ethernet, or MPLS at the access.

On the physical connection between the CE and PE, Frame Relay DLCI can still be used to indicate the particular service. However, these services are now carried through pre-established packet tunnels in the network. The provisioning is now simplified, because rather than configuring end-to-end PVCs in the network to establish connectivity between the different sites, the same can be accomplished by assigning the right DLCIs at the CE-to-PE connection; however, the services to different CEs are carried over pre-established tunnels in the network.

## L2VPN BGP Model

The L2VPN BGP model introduces some new terminology for referencing customer sites, customer equipment, and the way blocks of MPLS labels are allocated in the network.



**Figure 4-26** *Frame Relay Access and MPLS Edge/Core*

The L2VPN BGP model divides the network into two levels:

- **The provider backbone**—Contains all the PEs.
- **The sites**—These are the different locations where the customer equipment (CE) resides. A site can belong to a single customer and can have one or more CEs. Each CE is referenced using its own CE ID that is unique within the VPLS.

In other scenarios, a site can belong to multiple customers, in the case of an MTU, and each customer can have one or more CEs in that site. In this case, a customer connection is represented via a combination site ID and VPLS ID and a physical port on the MTU device.

In the BGP L2VPN scheme, each PE transmits pieces of information such as label blocks and information about the CEs to which it connects to all other PEs. To reach a destination, the PE need only install a route to the site where the destination exits. This allows the service to scale well, because this model tracks the number of VPN sites rather than individual customers.

The L2VPN BGP model is generalized to cover the following:

- **Connectivity of a CE to a PE**—In this model, the CE is a Frame Relay–capable or MPLS–capable device and can allocate a Frame Relay DLCI or an MPLS label after negotiation with the PE.
- **Connectivity of a CE to an L2PE and then to a PE**—This model reflects an MTU installation where an L2PE is used to connect the multiple customers within a site to a common piece of equipment in the basement. The L2PE is similar to the MTU-s that was already discussed. The L2PE is a switch that does MAC learning and bridging/switching,

JUNIPER Exhibit 1003

App. 6, pg. 126

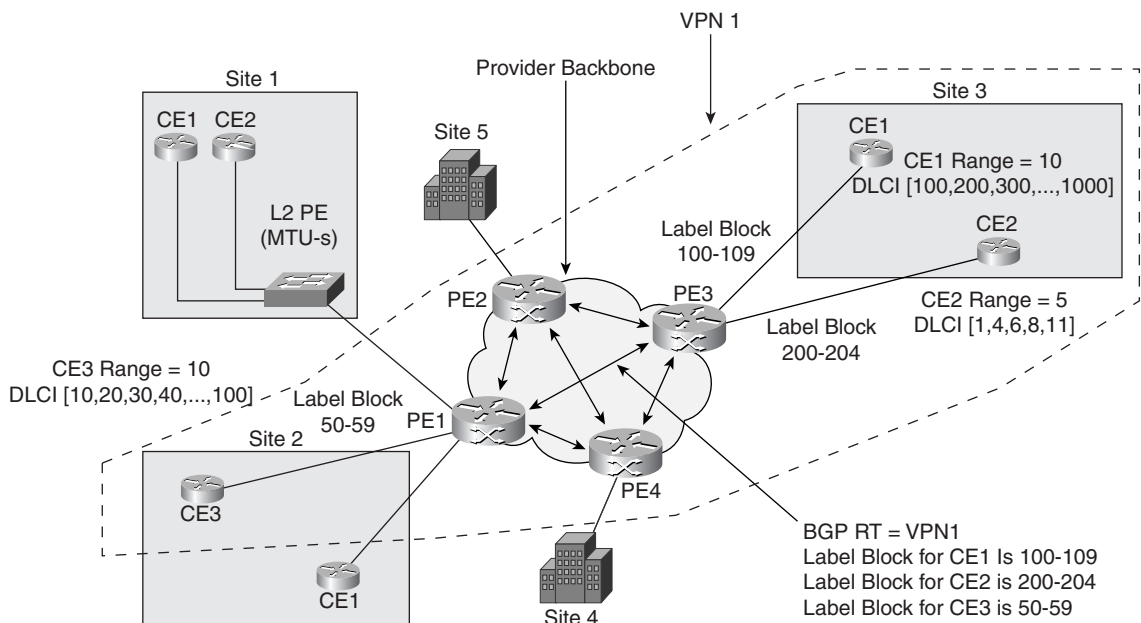
and it encapsulates the Ethernet customer traffic inside an MPLS packet with labels that are allocated by the PE. In the MTU case, the CE need only have an Ethernet connection to the L2PE and does not need to have any MPLS functionality.

The following section describes an example of an L2VPN with Frame Relay connectivity used on the access and MPLS used on the edge/core.

### Example of Frame Relay Access with MPLS Edge/Core

In this scenario, the CEs are connected to the PEs via Frame Relay, and the PEs carry the service over the network using MPLS. This is shown in Figure 4-27.

**Figure 4-27** *Frame Relay Access with MPLS Edge/Core*



Both CE and PE must agree on the FR DLCI that will be used on the interface connecting them. Each CE that belongs to a VPN is given a CE ID, which is unique in the context of the VPN. In Figure 4-27, a VPN consists of the three CEs: CE1, CE2, and CE3, where CE1 and CE2 are located in site 3 and CE3 is located in site 2. The CE IDs 1, 2, and 3 are supposed to be unique within the same VPN.

Each CE is configured with a list of Frame Relay DLCIs that allows it to connect to the other CEs in the VPN. The size of this list for a particular CE is called the CE's range. In Figure 4-27, for CE3 in site 2 to connect to both CE1 and CE2 in site 3, it would need two DLCIs, one for

each remote CE. As such, the CE range determines the number of remote sites a CE can connect to. The larger the range, the more remote CEs a CE can connect to. Each CE also knows which DLCI connects it to every other CE. When a packet comes to a CE from inside the customer network, the CE can use the correct DLCI based on where that packet is going. From then on, the packet is “switched” from one end to the other. The network behaves as a Frame Relay switch with respect to the CEs.

Each PE is configured with the VPNs in which it participates. For each VPN, the PE has a list of CEs that are members of that VPN. For each CE, the PE knows the CE ID, the CE range, and which DLCIs to expect from the CE. When a PE is configured with all the needed information for a CE, it chooses an MPLS label block, which is a contiguous set of labels. The number of these labels is the initial CE range, meaning if the CE has a range of ten DLCIs, the PE chooses ten MPLS labels. The smallest label in this label block is called the *label base*, and the number of labels in the label block is called a *label range*. The PE then uses BGP Network Layer Reachability Information (NLRI) to advertise the label blocks and the CE ID to which it connects, to all other PEs. Only the PEs that are part of the VPN (through the use of the BGP route target) accept the information. Other PEs discard the information or keep it for future use if they become part of that VPN.

A CE can have one or more label blocks, because when the VPN grows, more CEs participate in the VPN, and the CE label ranges might need to be expanded. If a CE has more than one label block, the notion of block offset is used. The *block offset* identifies the position of a label block in the set of label blocks of a given CE.

In reference to Figure 4-27, PEs 1, 2, and 3 participate in VPN1. The following is an example of the information that needs to be configured on PE1 and the information that PE1 advertises and learns via BGP.

Figure 4-27 shows the following:

- CE3 in site 2 and CE1 and CE2 in site 3 all belong to VPN1, as shown by the dotted line.
- CE3 is given the following set of DLCIs: [10,20,30,40,50,60,70,80,90,100], which correspond to a CE range of 10 (10 DLCIs).
- PE1 is given an MPLS label block that contains 10 labels, from label 50 to label 59. Label range = 10, label base = 50, block offset = 0.
- CE1 is given the following set of DLCIs: [100,200,300,400,500,600,700,800,900,1000], which correspond to a CE range of 10 (10 DLCIs).
- PE3's label block for CE1 contains 10 labels, from label 100 to label 109. Label range = 10, label base = 100, block offset = 0.
- CE2 is given the following set of DLCIs: [1,4,6,8,11], which correspond to a CE range of 5 (5 DLCIs).
- PE3's label block for CE2 contains 5 labels, from label 200 to label 204. Label range = 5, label base = 1, block offset = 0.

For PE1, the following takes place:

- 1 PE1 is configured as part of VPN1. BGP route target = VPN1.
- 2 PE1 is configured to have CE3 be part of VPN1. This can be done by configuring a physical port or a combination physical port and VLAN to be part of VPN1. CE1 is then assigned to that port/VLAN.
- 3 PE1 learns of CE1 and CE2 and the respective label blocks' offset and label base via BGP NLRI.
- 4 The following label information is configured on PE1 for CE3:
  - Label block: 50–59
  - Label base = 50
  - Label range (size of the block) = 10
  - Block offset = 0 (there is only one block)  
Note that PE1's label block is the same size as CE3's range of DLCIs, which is [10, 20, 30, 40, . . . , 100].
  - PE1 advertises the ID of CE3 and the label block to all other PEs via BGP NLRI.

The choice of assigning a DLCI to a particular CE is a local matter. Some simple algorithms could be used such that the CE ID of the remote CE becomes an index into the DLCI list of the local CE (with index 0 being the first entry in the list, 1 being the second entry in the list, and so on). So, for a connection between CE3 and CE1, CE3 could be allocated the second DLCI in the list (DLCI 20) because the remote CE is CE1, and CE1 is allocated the fourth DLCI in its list (DLCI 400) because the remote CE is CE3.

The PE in turn can use a simple algorithm to identify which MPLS label is used to reach a remote CE. In our example, suppose PE1 receives a BGP NLRI from PE3, indicating that CE1 has a label block 100–109. PE1 could use the CE ID of CE3 as an index into CE1's label block. In this case, PE1 could use label 103, which is CE1's label base (100) + CE3's ID (3). PE3 then uses label 51 (CE3's label base + 1) to reach CE3. As such, a packet coming from CE3 on DLCI 20 is encapsulated with MPLS label 103, and a packet coming from CE1 on DLCI 400 is encapsulated with MPLS label 51. An additional label is used on top of the stack to indicate the PE-to-PE LSP tunnel between PE1 and PE3.

## DTLS—Decoupling L2PE and PE Functionality

The concept discussed in the preceding section is extended to address the Ethernet-to-MPLS scenario—specifically, for MTU deployments. In an MTU scenario, multiple customers in the building are connected to a basement box, the MTU-s, which is referred to as “L2PE” in this section. In this case, the CEs are talking Ethernet to the L2PE, and the L2PE can talk either Ethernet or MPLS to the PE. Unlike the Frame Relay service, in which all the connections are P2P, the Ethernet service allows P2P and MP2MP VPLS service. In a multipoint service,

MAC addresses are used to distinguish how the traffic is directed over the MPLS network. MAC learning in PEs can cause scalability issues, depending on the L2 service. It is possible to decouple the functions needed to offer a VPLS service between the L2PE and the PE. These functions consist of the following:

- **MAC learning**—Learning MAC addresses from customers in the MTU and from other L2PEs across the metro
- **STP**—Building a loop-free topology on both the LAN side and the metro side
- **Discovery**—Discovering other L2PEs connected to the metro

It is possible to have the L2PE do the MAC learning and STP functions and to have the PE do the discovery function. As you know, the discovery function can be done via a protocol such as BGP to exchange information between the PEs. The benefit of this decoupling, called Decoupled Transparent LAN Service (DTLS), is to alleviate the PEs from the L2 functionality. Most PEs that have been deployed in provider networks are IP routers. These routers have been designed for IP core routing and L3 edge functionality and lack most of the functionality of L2 switches. L2 switches, on the other hand, come from an enterprise background and lack most of the scalability functions offered by L3 IP/MPLS routers.

Although new equipment is coming on the market that does both L2 switching and L3 IP routing, most of the deployed equipment are routers. DTLS allows the PE to function as IP routers and MPLS switches and puts all the L2 functionality in the L2PE. The PE does the VPN/VPLS discovery and runs BGP. The L2PE could then be a simpler L2 switch. The L2PE does not have to do any IP routing or run complex protocols such as BGP. The L2PE needs to be able to do L2 functions, such as MAC learning, STP, and bridging/switching, and be able to label the packets via either VLAN IDs or MPLS labels.

### Alleviating PEs from MAC Learning in the DTLS Model

One of the main issues for L2VPN services in the metro is MAC address learning. You have seen in this book that if the service offers a LAN connection between different sites and if the CEs are L2 switches (not routers), all MAC addresses that exist in the different connected LANs become part of the VPLS. As an example, assume the following scenario in a hypothetical metro:

- The metro contains 60 PEs
- Each PE is connected to ten buildings—that is, ten L2PEs
- Each L2PE services ten customers
- Each customer has two VPLS
- Each VPLS has 100 stations

The following calculations show the difference between starting the VPLS service at the L2PE and starting the service at the PE, based on the preceding information.

If you start the VPLS service at the L2PE, each L2PE has to support the following:

- Ten customers
- $10 * 2 = 20$  VPLS
- $20 * 100 = 2000$  MAC addresses

Also, because there are  $60 * 10 = 600$  L2PEs, assuming that each L2PE talks to every other L2PE in a full mesh (BGP or LDP), the number of bidirectional sessions between the L2PEs is  $600 * (600 - 1) / 2 = 179,700$  sessions.

Starting the VPLS service at the PE, the PE has to support the following:

- $10 * 10 = 100$  customers
- $100 * 2 = 200$  VPLS
- $200 * 100 = 20,000$  MAC addresses

If there is a full LDP/BGP mesh between the PEs, the number of bidirectional sessions is  $60 * (60 - 1) / 2 = 1770$  sessions.

From the previous calculations, the following can be easily deduced:

- Doing MAC learning at the L2PE and *not* at the PE scales much better. Otherwise, the PE has to deal with an explosion of MAC addresses.
- Doing a hierarchy in which the full mesh of BGP/LDP sessions starts at the PE prevents a session explosion.

The DTLS model keeps the MAC learning at the L2PE, assigns MPLS labels at the L2PE, and puts the VPLS discovery with BGP or other protocols at the shoulder of the PE. This way, the model can scale much better. The following needs to happen on the L2PE and on the PE:

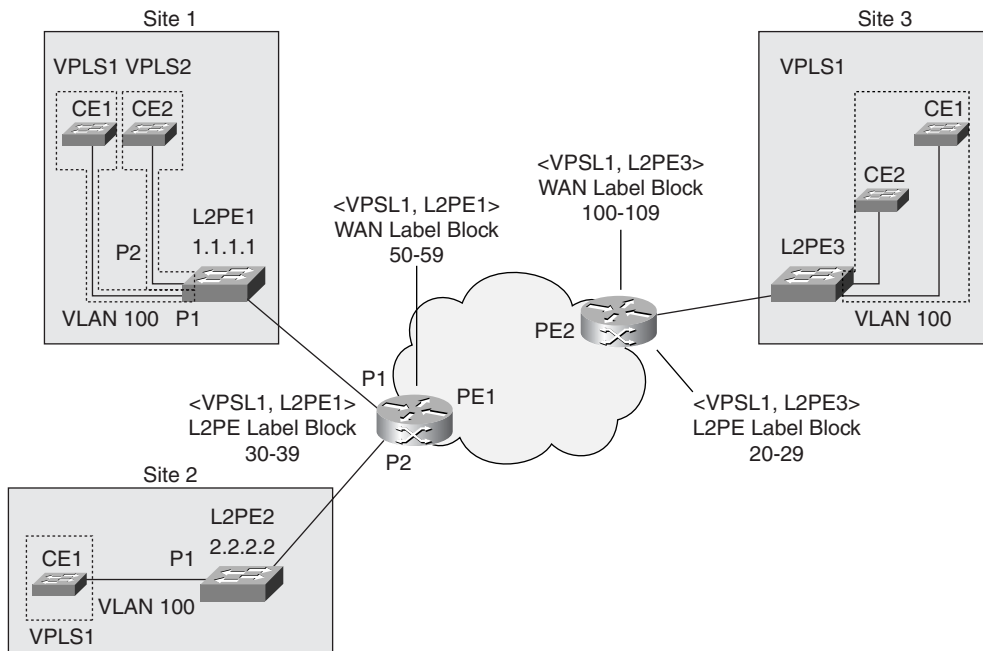
- The L2PE:
  - Needs to behave as a bridge/switch. It should be able to learn MAC addresses from the building customers and from other L2PEs.
  - Should be able to send or receive tagged packets. The L2PE should be able to perform tag stacking and swapping and handle both VLAN and MPLS tags.
  - Should be able to take an Ethernet frame from a customer-facing port (access port), strip the CRC and preamble, and encapsulate the remaining frame using an MPLS packet.
  - An L2PE that receives an MPLS packet should be able to decide which VPLS this packet belongs to and then send it to all customer-facing ports that belong to the VPLS.
  - Maintains mapping between the learned MAC addresses and the customer's ports and mapping between learned MAC addresses and labels. This mapping constitutes the L2PE's MAC address cache.
  - Maintains a separate MAC address cache per VPLS.

- Maintains a mapping between customer-facing ports and the different VPLS.
- Runs a protocol between it and the PE in a client/server model. The PE has the intelligence to discover the VPLSs in the network and to inform the L2PE of right labels (or label blocks, as described earlier). The L2PE uses these labels to reach its destination.
- The PE:
  - Needs to support the L2VPN functionality as described previously in the “L2VPN BGP Model” section. This means that a PE should be able to discover all the VPLSs in which it participates and distribute information about labels and about other L2PEs in the network.
  - Runs a PE-to-L2PE protocol that allows the decoupling of functionality between these two devices.

### Configuring the L2PE and PE

An L2PE needs to be told which VPLS it is a member of. This can be done by statically configuring a physical port or port/VLAN as part of a VPLS. In turn, for each (L2PE, VPLS) pair, the PE needs to be told the site ID of the (L2PE, VPLS). The PE also needs to be told which L2PEs it is connected to, and over which physical link and which VPLSs each L2PE participates in. This is illustrated in Figure 4-28.

**Figure 4-28** L2PE and PE Configuration



In Figure 4-28, L2PE1 is configured with the following:

- The L2PE1 ID is its router ID, 1.1.1.1
- Port 1 (P1) belongs to VPLS1
- Port 2 (P2) belongs to VPLS2
- L2PE1 is connected to PE1
- For the pair (L2PE1, VPLS1) the L2PE1 site ID is 1
- For the pair (L2PE1, VPLS2) the L2PE1 site ID is 1
- L2PE1 has a mapping between MAC addresses MAC x-MAC y with VPLS1
- L2PE1 has a mapping between MAC addresses MAC z-MAC w with VPLS2

If all information is configured on the PE, the PE can be given information pertinent to the L2PE that it can “push” into the L2PE via a certain client/server protocol. In this case, the PE needs to be configured with the following:

```
L2PE ID (router ID)
<connecting interface>
<VPLS ID> <L2PE site ID>
    <L2PE port ID, VLAN tag> <L2PE port ID, VLAN tag>
<VPLS ID> <L2PE site ID>
    <L2PE port ID, VLAN tag> <L2PE port ID, VLAN tag>
```

For each L2PE and each VPLS that the L2PE participates in, the PE is given the customer-facing port IDs and corresponding VLAN tags that belong to that VPLS. The PE then transfers all information relevant to that L2PE using the L2PE-PE protocol. The protocol that allows the information exchange between the L2PE and PE can be an extension to LDP or via other protocols. The PE transfers all information relevant to other PEs using the L2VPN BGP mechanism.

The following is a sample configuration for PE1, as shown in Figure 4-28. In this example, PE1 is connected to two different sites, 1 and 2. In site 1, L2PE1 offers service to two customers. Customer 1 on port 1 has VPLS1, which emulates a LAN between VLAN 100 across different sites (sites 1, 2, and 3). Customer 2 has VPLS2, which emulates a LAN for all VLANs.

PE1 is connected to L2PE1 and L2PE2. The following is the configuration for L2PE1 in PE1:

- L2PE1 has router ID 1.1.1.1
- Connecting interface: P1
- <VPLS 1, site ID 1>
  - <L2PE1 port 1, VLAN 100>
- <VPLS 2, site ID 1>
  - <L2PE1 port 2, all>



- For VPLS1:
  - WAN label block 50–59, label range = 10, label base = 50, block offset = 0
  - L2PE label block 30–39, label range = 10, label base = 30, block offset = 0
- For VPLS2:
  - WAN label block is x
  - L2PE label block is y

The following is the configuration for L2PE2 in PE1:

- L2PE2 has router ID 2.2.2.2
- Connecting interface: P2
- <VPLS 1, site ID 2>
  - <L2PE2 port 1, VLAN 100>
- For VPLS1:
  - WAN label is etc.
  - L2PE label block is etc.

Note that the PE1 configuration includes the indication of label blocks and label ranges. This is the same concept discussed earlier for the Frame Relay scenario; however, two sets of label blocks need to be configured for each PE. One set, called the *WAN label block*, is used to direct traffic received from L2PEs served by other PEs to the correct L2PE served by this PE. The other set of label blocks is the L2PE label block that tells the L2PEs which label to use when sending traffic to another L2PE. This creates in the network a hierarchy where the L2PEs exchange information with the connected PEs and the PEs exchange information with each other.

The site ID of an L2PE could be used as an offset from a label base to create a label. The next two sections explain how this is applied for WAN labels and the L2PE labels.

## WAN Labels

For VPLS1, PE3, which is connected to L2PE3 in site 3, sends a BGP advertisement to PE1. This advertisement contains PE3's WAN label base of 100, a block offset of 0 (because only one label block is used), and the label range of 10.

For VPLS1, PE1, which is connected to L2PE1 in site 1, sends a BGP advertisement to PE3. This advertisement contains PE1's WAN label base of 50, a block offset of 0, and a label range of 10. PE1 also sends advertisements for all the <L2PE, VPLS> pairs it connects to, such as <L2PE1, VPLS2> and L2PE2 and its respective VPLS.

PE1 uses label 101 when sending packets to L2PE3. This is calculated by taking PE3's label base (100) and adding PE1's site ID (1). PE3 uses label 53 when sending packets to L2PE1. This is calculated by taking PE1's label base (50) and adding PE3's site ID (3).

## L2PE Labels

Each PE also allocates a set of label blocks, called *L2PE labels*, that will be used by the L2PEs. For VPLS1, PE1 sends to L2PE1 a label base of 30, a label range of 10, and a block offset of 0. For VPLS1, PE3 sends to L2PE3 a label base of 20 and a range of 10.

L2PE1 uses label 33 when sending packets to PE1. This is calculated by taking PE1's label base (30) and adding L2PE3's site ID (3). L2PE3 uses label 21 when sending packets to PE3. This is calculated by adding PE3's label base (20) to L2PE1's site ID (1).

Following a packet in VPLS1 from L2PE1, L2PE1 takes the Ethernet frame coming from port 1, VLAN 100, and encapsulates it in an MPLS frame with label 33. PE1 receives the packet with label 33 and swaps this label with label 101, which is sent to L2PE3. PE1 encapsulates another PE-to-PE label, which directs the packet from PE1 to PE3. When the packet reaches PE3, PE3 swaps the label 103 for a label 21 and directs the packet to L2PE3. Based on this label, L2PE3 directs the packet to VPLS1.

Flooding, learning, and spanning-tree behavior at the L2PE are similar to what was previously described with the L2VPN and the LDP PW model. When a packet with an unknown destination reaches the L2PE, the L2PE identifies to which VPLS this packet belongs. It then replicates the packet over all ports in the VPLS. If the packet is received on a customer-facing port, the L2PE sends a copy out every other physical port or VLAN that participates in the VPLS, as well as to every other L2PE participating in the VPLS. If the packet is received from a PE, the packet is sent to only customer-facing ports in the MPLS, assuming that a full mesh of PEs already exists.

If an L2PE wants to flood a VPLS packet to all other L2PEs in the VPLS, the L2PE sends a copy of the packet with each label in the L2PE label ranges for that VPLS, except for the label that corresponds to the L2PE itself.

The drawback of doing the flooding at the L2PE is that the L2PE is connected to many other L2PEs in other sites and has to do quite a lot of replications. You have to weigh this against the benefits of removing the MAC learning from the PEs and keeping it in the L2PEs.



As mentioned, the protocol used for the PE-to-L2PE information exchange can be an extension of LDP. Also, there is no technical restriction on whether the tags used between the L2PE and the PE are MPLS labels. Using VLAN tags with Q-in-Q is also a possibility. The choice of one approach or the other is implementation-specific and depends on the L2PE and PE equipment capability. The upper VLAN tag sent between the L2PE and the PE is indicative of the VPLS. The PE needs to match that tag with the right WAN label to transport the packet to the remote L2PEs. It is also possible to use LDP as a universal protocol to allow the exchange of Q-in-Q tags between the PE and L2PE in the same way that MPLS labels are exchanged.

## Conclusion

You have seen in this chapter how IP/MPLS can be used to scale L2 Ethernet service deployments. By keeping L2 Ethernet networks confined to the access/edge and IP/MPLS at the edge/core, service providers can leverage the simplicity of deploying Ethernet LANs with the scalability offered by IP and MPLS. L2 Ethernet services can be offered as P2P or MP2MP services. P2P can be achieved via mechanisms such as L2TPv3 or EoMPLS draft-martini. MP2MP can be achieved via VPLS.

You have seen that the flexibility VPLS offers with any-to-any connectivity is also coupled with the drawbacks of delivering Ethernet LANs in dealing with L2 loops and broadcast storms. Also with VPLS come the challenges of dealing with MAC address explosion, because PEs have to keep track of all MAC addresses advertised within the VPLS(s) the PEs belong to. Some alternatives, such as DTLS, are proposed for dealing with MAC explosion; however, different network designs and different L2PE-to-PE protocols would have to be defined and standardized.

Part II of this book, starting with Chapters 5 and 6, builds on the fact that scalable L2VPN networks are built with hybrid Ethernet and IP/MPLS networks. It also focuses on scaling the MPLS portion of the network with mechanisms such as traffic engineering via RSVP-TE and traffic protection via MPLS fast reroute. Chapters 7 and 8 move into the more advanced topic of Generalized MPLS (GMPLS). Metro networks are built with legacy TDM technology, so it is important to understand how the proliferation of MPLS in the metro will affect network provisioning on both packet and TDM networks—hence the need for a generalized control plane like GMPLS.





JUNIPER Exhibit 1003  
App. 6, pg. 137

From the Library of Tal Lavian



# MPLS: Controlling Traffic over Your Optical Metro

- Chapter 5 MPLS Traffic Engineering
- Chapter 6 RSVP for Traffic Engineering and Fast Reroute
- Chapter 7 MPLS Controlling Optical Switches
- Chapter 8 GMPLS Architecture



This chapter covers the following topics:

- Advantages of Traffic Engineering
- Pre-MPLS Traffic Engineering Techniques
- MPLS and Traffic Engineering

# MPLS Traffic Engineering

---

You have seen in the previous chapters how metro Ethernet Layer 2 (L2) services can be deployed over an MPLS network. You also learned about the concept of pseudowires and label switched path (LSP) tunnels. The LSP tunnels are simply a means to tunnel the pseudowires from one end of the MPLS cloud to the other with the opportunity of aggregating multiple pseudowires within a single LSP tunnel. The LSP tunnels themselves can be constructed manually, or via MPLS signaling using the Label Distribution Protocol (LDP) or RSVP traffic engineering (TE). TE is an important MPLS function that gives the network operator more control over how traffic traverses the network. This chapter details the concept of TE and its use.

## Advantages of Traffic Engineering

One of the main applications of MPLS is TE. A major goal of Internet TE is to facilitate efficient and reliable network operations while simultaneously optimizing network resource utilization and traffic performance. TE has become an indispensable function in many large provider networks because of the high cost of network assets and the commercial and competitive nature of the Internet.

The purpose of TE is to optimize the performance of operational networks. TE forces packets to take predetermined paths to meet network policies. In general, TE provides more efficient use of available network resources; provides control of how traffic is rerouted in the case of failure; enhances performance characteristics of the network relative to packet loss, delay, and so on; and enables value-added services, such as guaranteeing QoS and enforcing SLAs.

With metro Ethernet services, you have seen that setting bandwidth parameters on the UNI connection between the customer edge (CE) and the provider edge (PE) devices is part of the service sold to the customer. An Ethernet service with a committed information rate (CIR) of 1 Mbps should guarantee the customer that much bandwidth. It is the service provider's duty to make sure that the bandwidth promised to the customer can be allocated on the network and that the traffic adheres to the packet loss and delay parameters that are promised. TE gives the service provider more control over how traffic from multiple customers is sent over the network, enabling the service provider to make the most use of the resources available and to optimize performance.

**JUNIPER Exhibit 1003**  
**App. 6, pg. 140**

In reference to RFC 2702, *Requirements for Traffic Engineering over MPLS*, the key performance objectives for TE can be classified as either of the following:

- Traffic-oriented
- Resource-oriented

Traffic-oriented performance objectives deal with traffic characteristics such as minimizing loss and delay to enhance traffic quality. In reference to the performance parameters defined in Chapter 3, “Metro Ethernet Services,” traffic characteristics include availability, delay, jitter, and packet loss.

Resource-oriented performance objectives are mainly concerned with the optimization of resource utilization. The top priority of these objectives is to manage bandwidth resources through congestion control. Network congestion typically manifests under two scenarios:

- When network resources are insufficient or inadequate to accommodate the traffic load. An example is a spoke between a multitenant unit (MTU) device and a provider edge router at the central office (CO) that has less bandwidth than required to service all the customers of the building according to an agreed-upon SLA with the service provider.
- When traffic streams are inefficiently mapped onto available resources, causing subsets of network resources to become overutilized while others remain underutilized. An example is the existence of multiple parallel links on the backbone where some of these links are oversubscribed and are dropping traffic while others are sitting idle. This is because of how Interior Gateway Protocols (IGPs) calculate the shortest path, as explained in the next section.

Expanding capacity, or overprovisioning, alleviates the first type of congestion. Adding more or bigger network pipes is a quick and easy fix, but it comes at additional cost. Other classic congestion-control techniques, such as rate limiting, queue management, and others, can also be used to deal with insufficient network resources. These techniques are important to prevent a set of users or traffic types from consuming the whole bandwidth and starving other users on the network.

This chapter mainly addresses the second type of congestion problems—those resulting from inefficient resource allocation. You can usually address these congestion problems through TE. In general, you can reduce congestion resulting from inefficient resource allocation by adopting load-balancing policies. The objective of such strategies is to minimize maximum congestion, or alternatively to minimize maximum resource utilization, through efficient resource allocation. When congestion is minimized through efficient resource allocation, packet loss decreases, transit delay decreases, and aggregate throughput increases. This significantly enhances the end users’ perception of network service quality.

As you have noticed, this chapter so far hasn’t mentioned MPLS, because TE by itself is universal and a well-understood problem. The use of MPLS for TE is one of the methods for dealing with resource optimization, and the industry has begun adopting MPLS techniques only after going through many alternatives to solve the TE problem. The next section discusses some of the pre-MPLS TE techniques.



## Pre-MPLS Traffic Engineering Techniques

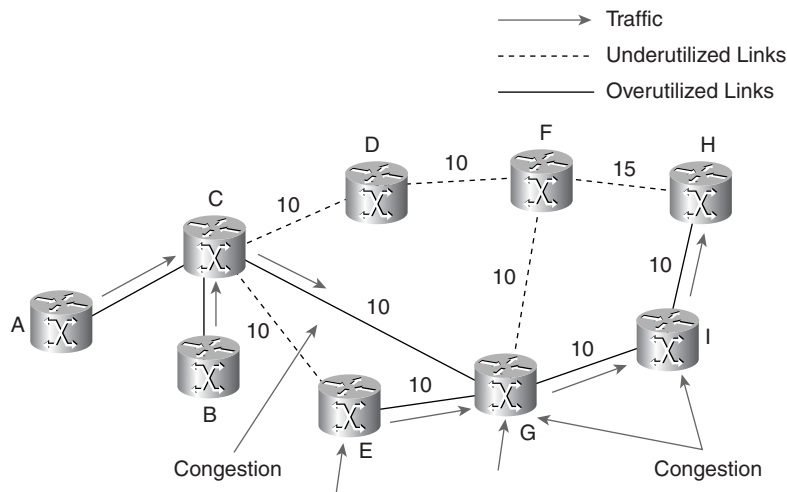
Pre-MPLS TE techniques involved multiple mechanisms:

- Altering IGP routing metrics
- Equal-cost multipath
- Policy-based routing
- Offline design of virtual circuit overlays

### Altering IGP Routing Metrics

IGPs have many limitations when used to achieve traffic engineering. IGPs rely on metrics that do not reflect actual network resources and constraints. IGPs based on Shortest Path First (SPF) algorithms contribute significantly to congestion problems in IP networks. SPF algorithms generally optimize based on a simple additive metric. These protocols are topology-driven, so real-time bandwidth availability and traffic characteristics are not factors considered in routing decisions. As such, congestion occurs when the shortest paths of multiple streams converge over one link that becomes overutilized while other existing links are underutilized, as shown in Figure 5-1.

**Figure 5-1** IGP Shortest Path First Congestion



In Figure 5-1, based on the indicated link metric, an OSPF routing algorithm allows traffic coming from routers A and B, destined for router H, to use path C-G-I-H. Traffic from routers E and G, destined for H, uses path G-I-H. As you can see, multiple streams of traffic have converged on the same links or routers, causing congestion on link G-I, for example, while other links in the network remain underutilized.

Altering IGP metrics could cause traffic to shift between links. Changing the metric of link F-H from 15 to 10 or 5 could cause the traffic to start taking links C-D-F-H or C-G-F-H. Link manipulation for the purposes of TE works for quick-fix solutions but is mainly a trial-and-error process and does not scale as an operational model. Adjusting the link metrics might fix one problem but create other problems in the network.

## Equal-Cost Multipath

Equal-cost multipath is a mechanism that allows routers to distribute traffic between equal-cost links to efficiently use the network resources and avoid the problem of link oversubscription and undersubscription. If, for example, a router calculates the shortest path based on link metrics and determines multiple equal paths exist to the same destination, the router can use load-balancing techniques to spread the traffic flows over the equal-cost links. Referring to Figure 5-1, if the metric of link F-H is changed to 10 instead of 15, the paths C-D-F-H and C-G-I-H would have the same metric, 30 (10 + 10 + 10). Traffic from routers A and B, destined for H, could be load balanced across the two equal-cost paths.

## Policy-Based Routing

Policy-based routing is another mechanism that can be used for TE. It allows routers to dictate the traffic trajectory. That is, they pick the router output interface on which to route traffic based on a policy—for example, based on the source IP address rather than the destination IP address. With this type of TE, you can dictate that traffic coming from a certain customer or provider goes one way, while traffic from other customers or providers goes the other way, irrespective of what its actual destination is.

Policy-based mechanisms can be used to allow more granularity in identifying the source traffic. For example, the traffic can be identified based on source IP address, router port numbers, QoS, application type, and so on. Although this type of TE is useful, it has its drawbacks. First, it acts against the nature of routing, which is primarily destination-based. Second, it becomes yet another form of intelligent static routing with vulnerability to traffic loops and to the lack of dynamic rerouting in case of failure of network elements.

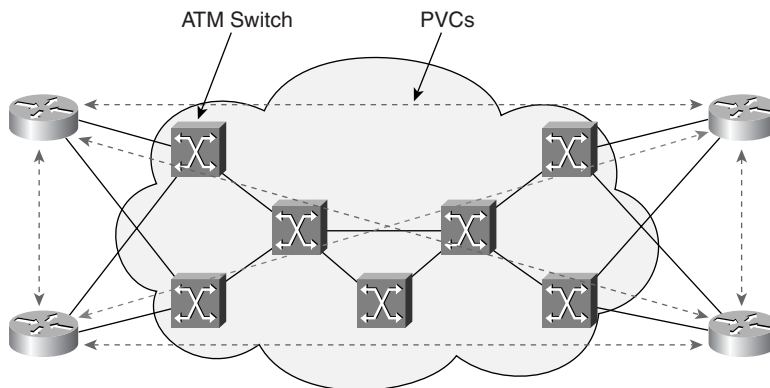
## Offline Design of Virtual Circuit Overlays

A popular approach to circumvent the inadequacies of current IGPs is to use an overlay model, such as IP over ATM or IP over Frame Relay. The overlay model extends the design space by enabling arbitrary virtual topologies to be provisioned on top of the network's physical topology. The virtual topology is constructed from virtual circuits (VCs) that appear as physical links to the IGP routing protocols. Overlay techniques can range from simple permanent virtual circuit (PVC) provisioning between routed edge networks to more fancy mechanisms

that include constraint-based routing at the VC level with support of configurable explicit VC paths, traffic shaping and policing, survivability of VCs, and so on.

Figure 5-2 shows edge routers that are connected to each other via an overlay model on top of an ATM network. For the IGPs, the VCs appear as direct physical links between the routers. Traffic can be engineered between the routed edges and is agnostic to the L2 switched network in the middle of the cloud. This type of TE has several benefits: It enables you to achieve full traffic control, measure link statistics, divert traffic based on link utilization, apply load-balancing techniques, and so on. It also has several obvious drawbacks: It creates multiple independent control planes—IP and ATM—that act independently of one another, a full mesh between the routers, an IGP neighbor explosion (each router has all other routers as neighbors and has to exchange routing updates with them), and finally a network management challenge constituting multiple control layers.

**Figure 5-2** IGP TE Via Virtual Circuit Overlays



## MPLS and Traffic Engineering

MPLS is strategically significant for TE because it can potentially provide most of the functionality available from the overlay model (described in the preceding section), with much better integration with IP. MPLS for TE is attractive because it enables you to do the following:

- Manually or dynamically build explicit LSPs
- Efficiently manage LSPs
- Define traffic trunks and map them to LSPs
- Associate a set of attributes with traffic trunks to change their characteristics
- Associate a set of attributes with resources that constrain the placement of LSPs and traffic trunks mapped to those LSPs
- Aggregate and deaggregate traffic (whereas IP routing only allows aggregation)

JUNIPER Exhibit 1003

App. 6, pg. 144

- Easily incorporate a constraint-based routing framework with MPLS
- Deliver good traffic implementation with less overhead than pre-MPLS techniques
- Define backup paths with fast failover

Before delving into more details about TE, it helps to explain the terminology of trunks versus LSPs, because the two are often confused with one another.

## Traffic Trunks Versus LSPs

Traffic trunks are not LSPs. The definition of traffic trunks as indicated in RFC 2430 follows: “A traffic trunk is an aggregation of traffic flows of the same class which are placed inside an LSP.” Examples of flow classes can be similar to Diffserv. Traffic trunks are also routable objects, similar to VCs for ATM. A traffic trunk can be mapped to a set of LSPs and can be moved from one LSP to another.

An LSP, on the other hand, is a specification of the path through which the traffic traverses. The LSP is constructed through label swapping between ingress to egress to switch the traffic to its destination. Trunks traverse LSPs and can be routed from one LSP to another. This is illustrated in Figure 5-3.

**Figure 5-3** *Trunks and LSPs*

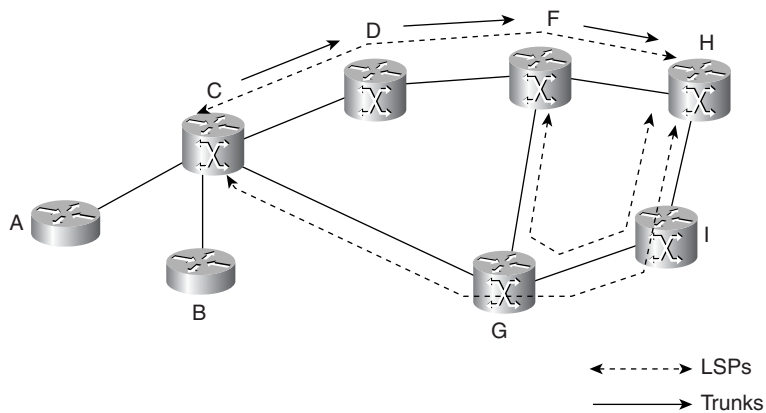


Figure 5-3 shows two LSPs between routers C and H, LSP C-D-F-H and LSP C-G-I-H. Another LSP exists between routers F and H, LSP F-G-I-H. A set of traffic flows belonging to the same class, coming from router A and destined for destinations beyond router H, could be mapped to either LSP C-D-F-H or LSP C-G-I-H. This aggregated traffic flow is the traffic trunk. The same traffic trunk can be routed over LSP F-G-I-H if some trunk attributes, such as resiliency or bandwidth, are being enforced.

## Capabilities of Traffic Engineering over MPLS

The functional capabilities required to support TE over MPLS in large networks involve the following:

- A set of attributes that affect the behavior and characteristics of traffic trunks
- A set of attributes that are associated with resources and that constrain the placement of traffic trunks over LSPs
- A constraint-based routing framework that is used to select paths subject to constraints imposed by traffic trunk attributes and available resources

The attributes associated with traffic trunks and resources, as well as parameters associated with routing, represent a set of variables that can be used to engineer the network. These attributes can be set either manually or through automated means. The next section discusses traffic trunk operation and attributes.

### Traffic Trunk Operation and Attributes

Traffic trunks are by definition unidirectional, but it is possible to instantiate two trunks in opposite directions with the same endpoints. The set of traffic trunks, one called *forward trunk* and the other called *backward trunk*, form a logical bidirectional traffic trunk. The bidirectional traffic trunks can be topologically symmetrical or asymmetrical. A bidirectional traffic trunk is symmetrical if opposite trunks take the same physical path, and it is asymmetrical if opposite trunks take different physical paths.

The basic operations that you can perform on a trunk include establishing a trunk; activating, deactivating, and destroying a trunk; modifying a trunk's attributes; and causing a trunk to reroute from its original path via manual or dynamic configuration. You can also police the traffic to comply with a certain SLA and shape and smooth the traffic before it enters the network.

As described in RFC 2702, the following are the basic attributes of traffic trunks that are particularly significant for TE:

- Traffic parameter attributes
- Generic path selection and maintenance attributes
- Priority attribute
- Preemption attribute
- Resilience attribute
- Policing attribute
- Resource attributes

### Traffic Parameter Attributes

Traffic parameter attributes indicate the resource requirements of a traffic trunk that are useful for resource allocation and congestion avoidance. Such attributes include peak rates, average rates, permissible burst size, and so on. Chapter 3 describes the applicable parameters, such as committed information rate (CIR), peak information rate (PIR), and so on.

### Generic Path Selection and Maintenance Attributes

Generic path selection and maintenance attributes define how paths are selected, such as via underlying network protocols, via manual means, or via signaling. If no restrictions exist on how a traffic trunk is established, IGPs can be used to select a path. If restrictions exist, constraint-based routing signaling, such as RSVP-TE, should be used.

Chapter 4, “Hybrid L2 and L3 IP/MPLS Networks,” describes how a metro provider carries L2 services over an MPLS cloud via the use of pseudowires (VC-LSPs) carried in LSP tunnels. If the LSP tunnels are not traffic-engineered, the traffic on the MPLS cloud follows the path dictated by the IGP. If multiple IGP paths collide, traffic congestion could occur. Setting resource requirements coupled with TE alleviates this problem.

### Priority Attribute

The priority attribute defines the relative importance of traffic trunks. Priorities determine which paths should be used versus other paths at connection establishment and under fault scenarios. A metro operator could deliver Internet service as well as IP storage backhaul over different pseudowires. The IP storage traffic could be carried over a separate LSP tunnel and given a high priority to be rerouted first in case of failure.

### Preemption Attribute

The preemption attribute determines whether a traffic trunk can preempt another traffic trunk from a given path. Preemption can be used to ensure that high-priority traffic can always be routed in favor of lower-priority traffic that can be preempted. Service providers can use this attribute to offer varying levels of service. A service that has preemption could be priced at a higher rate than a regular service.

### Resilience Attribute

The resilience attribute determines the behavior of a traffic trunk when fault conditions occur along the path through which the traffic trunk traverses. The resiliency attribute indicates whether to reroute or leave the traffic trunk as is under a failure condition. More extended resilience attributes could specify detailed actions to be taken under failure, such as the use of alternate paths, and specify the rules that govern the selection of these paths.

## Policing Attribute

The policing attribute determines the actions that should be taken by the underlying protocols when a traffic trunk exceeds its contract as specified in the traffic parameters. Policing is usually done on the input of the network, and it indicates whether traffic that does not conform to a certain SLA should be passed, rate limited, dropped, or marked for further action.

## Resource Attributes

Resource attributes constrain the placement of traffic trunks. An example of resource attributes is the maximum allocation multiplier. This attribute applies to bandwidth that can be oversubscribed or undersubscribed. This attribute is comparable to the subscription and booking factors in ATM and Frame Relay. A resource is overallocated or overbooked if the sum of traffic from all traffic trunks using that resource exceeds the resource capacity. Overbooking is a typical mechanism used by service providers to take advantage of the traffic's statistical multiplexing and the fact that peak demand periods for different traffic trunks do not coincide in time.

Another example of resource attributes is the resource class attribute, which attempts to give a "class" to a set of resources. Resource class attributes can be viewed as "colors" assigned to resources such that resources with the same "color" conceptually belong to the same class. The resource class attribute can be used to implement many policies with regard to both traffic- and resource-oriented performance optimization. Resource class attributes can be used, for example, to implement generalized inclusion and exclusion to restrict the placement of traffic trunks to a specific subset of resources.

## Constraint-Based Routing

Constraint-based routing assists in performance optimization of operational networks by finding a traffic path that meets certain constraints. Constraint-based routing is a demand-driven, reservation-aware routing mechanism that coexists with hop-by-hop IGP routing.

Constraints are none other than the attributes that were previously discussed: Trunk attributes such as path selection attributes, policing, preemption, and so on, coupled with resource attributes and some link-state parameter, would affect the characteristics and behavior of the traffic trunk.

A constraint-based routing framework can greatly reduce the level of manual configuration and intervention required to set TE policies. In practice, an operator specifies the endpoints of a traffic trunk and assigns a set of attributes to the trunk. The constraint-based routing framework is then expected to find a feasible path to satisfy the expectations. If necessary, the operator or a TE support system can then use administratively configured explicit routes to perform fine-grained optimization.

Figures 5-4a and 5-4b show two different types of routing applied to the same scenario. In Figure 5-4a, simple IGP routing is applied, and the shortest path is calculated based on the IGP

metrics. A traffic trunk coming from router A is mapped to path (LSP) C-E-G-I-H. In Figure 5-4b, constraints are imposed on the routing construct. The constraint is *not* to use any path that has available bandwidth less than 250 Mbps. As such, the two links E-G and G-I have been removed, or pruned, from the selection algorithm, and the traffic trunk coming from router A has been mapped to path C-D-G-F-H.

Figure 5-4a Aggregating Trunks into Tunnels

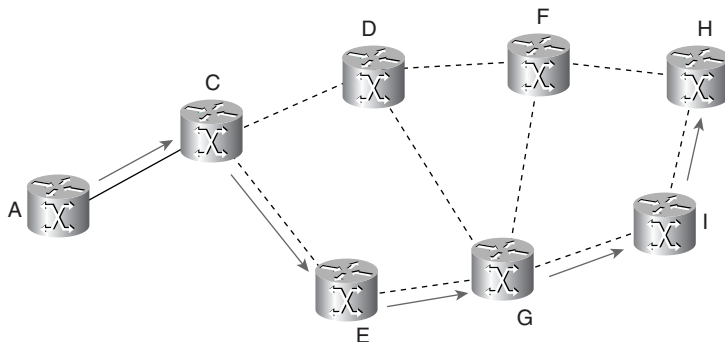
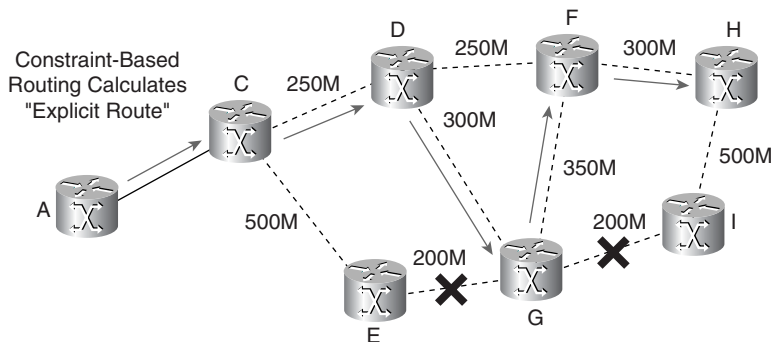


Figure 5-4b Constraint-Based Routing



## Conclusion

This chapter has discussed the different parameters used for TE. Some of the concepts, such as traffic parameter attributes and policing attributes, were discussed in the context of metro deployments in Chapter 3.

The next steps for TE entail a mechanism for exchanging the traffic attributes and parameters in the network for each router to build a TE database. This database gives the routers visibility



to all the network resources and attributes that can be used as input into a Constrained Shortest Path First (CSPF) algorithm. CSPF determines the path in the network based on different constraints and attributes. Finally, a signaling protocol such as RSVP-TE is used to signal the LSP in the network based on the path determined by the CSPF. The next chapters explain the concepts behind RSVP-TE to familiarize you with how Label Switched Path are signaled across a packet network. The book extends this concept further in Chapters 7 and 8 to discuss how MPLS signaling and routing can be extended as well to cover nonpacket networks, such as the case of an optical metro.



This chapter covers the following topics:

- Understanding RSVP-TE
- Understanding MPLS Fast Reroute

# RSVP for Traffic Engineering and Fast Reroute

---

Traffic engineering allows the service provider to manipulate the traffic trajectory to map traffic demand to network resources. You have seen in Chapter 5, “MPLS Traffic Engineering,” that traffic engineering can be achieved by manipulating Interior Gateway Protocol (IGP) metrics or, better yet, by using a signaling protocol such as RSVP-TE. RSVP-TE offers the ability to move trunks away from the path selected by the ISP’s IGP and onto a different path. This allows a network operator to route traffic around known points of congestion in the network, thereby making more efficient use of the available bandwidth. It also allows trunks to be routed across engineered paths that provide guaranteed service levels, enabling the sale of classes of service.

In metro networks, traffic engineering goes hand in hand with traffic path restoration upon failures. The behavior of a network upon failure depends on what layer the restoration methods are applied to. SONET/SDH networks, for example, can achieve restoration at Layer 1, meaning that if part of a SONET/SDH ring fails, there is always a backup TDM circuit provisioned on another fiber (unidirectional path switched ring, UPSR) or another pair of fibers (bidirectional line switched ring, BLSR). With Resilient Packet Ring (RPR), the ring is always fully utilized and a failure will cause the ring to wrap, allowing the rest of the ring to remain functional.

Restoration can also be done at Layer 2. Spanning Tree Protocol (STP) and Rapid Spanning Tree Protocol (RSTP) (802.1w) are typical methods that allow the network to converge after failure. Layer 3 methods can also be used. Routing protocols such as Open Shortest Path First (OSPF) and Intermediate System-to-Intermediate System (IS-IS) are capable of computing multiple paths to the same destination. If the main path fails, the protocols converge to an alternate path. Mechanisms like equal-cost multipaths can also be used to allow faster convergence by having parallel active paths to the same destination.

In the use of traffic engineering and traffic restoration methods, operators look for the following:

- The ability to maintain customer SLAs in case of a network failure.
- The ability to achieve the most efficient use of network resources, in a way that provides good QoS that meets an SLA with their customers.
- The ability to restore failure within a timeline that does not violate any SLAs they established with their customers.

Because MPLS plays a big role in delivering and scaling services in the metro, it is important to understand how it can be used to achieve TE and protection via the use of RSVP-TE. In this chapter, you see how MPLS, through the use of RSVP-TE, can be used to establish backup paths in the case of failure.

## Understanding RSVP-TE

MPLS TE may be used to divert traffic over an explicit route. The specification of the explicit route is done by enumerating an explicit list of the routers in the path. Given this list, TE trunks can be constructed in a variety of ways. For example, a trunk could be manually configured along the explicit path. This involves configuring each router along the path with state information for forwarding the particular MPLS label.

Alternately, a signaling protocol such RSVP-TE can be used with an EXPLICIT\_ROUTE object (ERO) so that the first router in the path can establish the trunk. The ERO is basically a list of router IP addresses.

### NOTE

Constraint-based routing LDP (CR-LDP) is another signaling protocol that can be used to build traffic-engineered paths. However, the use of RSVP-TE is more widely deployed and as such will be discussed in this book.

Originally, RSVP (defined in RFC 2205, *Resource ReSerVation Protocol—Version 1 Functional Specification*) was designed as a protocol to deliver QoS in the network by allowing routers to establish resource reservation state for individual flows originated between hosts (computers). This model has not taken off with network operators because of scalability issues in maintaining the per-flow state between pairs of hosts in each router along the IGP path. The RSVP implementation is illustrated in Figure 6-1.

**Figure 6-1** Original RSVP Implementation

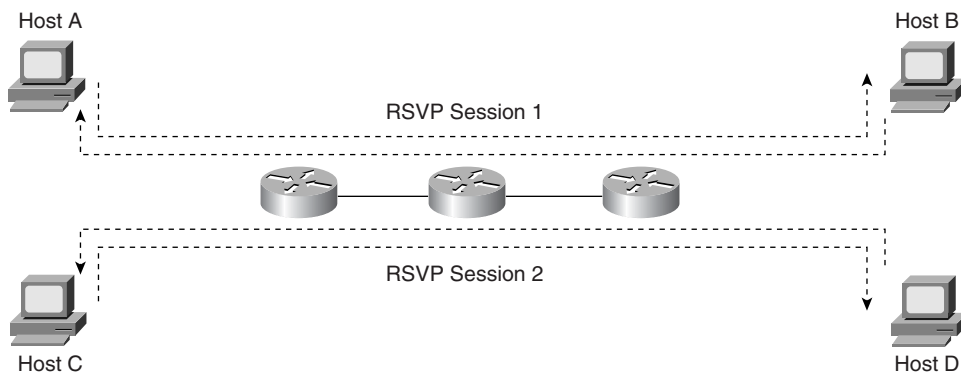
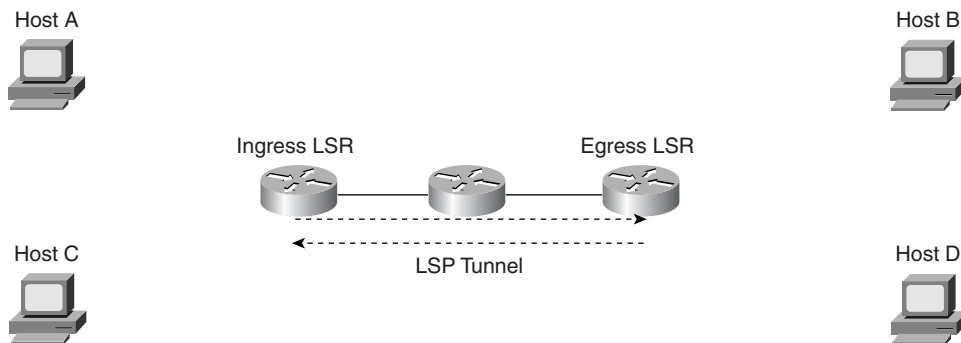


Figure 6-1 illustrates two RSVP sessions between hosts A and B and hosts C and D. The routers in the path would have to maintain state information for these sessions to allocate certain bandwidth to the individual flows. With a large number of hosts (millions) in a public network, this model has not proven to be efficient and hence has not been adopted in the public Internet.

In the late 1990s, RSVP was extended to support the creation of MPLS label switched paths (LSPs). The extended RSVP implementations introduced a lot of changes to the traditional RSVP, to support scalability issues and TE. In particular, RSVP sessions take place between ingress and egress label switch routers (LSRs) rather than individual hosts. The aggregated traffic flow, called a *traffic trunk*, is then mapped to LSPs, also called *LSP tunnels*. The RSVP-TE implementation is shown in Figure 6-2.

**Figure 6-2** *RSVP-TE Implementation*



The extensions of RSVP to support MPLS and TE can accomplish the following:

- **Establish a forwarding path**—RSVP can be used to establish LSPs by exchanging label information. This mechanism is similar to the Label Distribution Protocol (LDP).
- **Establish an explicit path**—RSVP-TE is used to establish an LSP along an explicit route according to specific constraints. LSPs can be rerouted upon failure. (Fast reroute is discussed later, in the section “Understanding MPLS Fast Reroute.”)
- **Resource reservation**—The existing RSVP procedures for resource reservation can be applied on aggregated flows or traffic trunks. This model scales because it is done on trunks rather than flows and is done between pairs of routers rather than pairs of hosts, as was originally intended for RSVP.

The reason IETF chose to extend RSVP to support MPLS and TE has to do with the fact that RSVP was originally designed for resource reservation in the Internet, a concept that is closely tied to TE, so it makes sense to extend the protocol rather than create a new one. RSVP also can carry opaque objects such as fields that can be delivered to routers, which makes it easy to define new objects for different purposes. The purpose of some of these objects is to carry labels for the label distribution function, whereas the purpose of others is to create explicit routes.

The following sections describe how RSVP tunnels are created and the mechanisms that are used to exchange MPLS labels and reserve bandwidth:

- RSVP LSP Tunnels
- Label Binding and LSP Tunnel Establishment Via RSVP
- Reservation Styles
- Details of the PATH Message
- Details of the Reservation Message

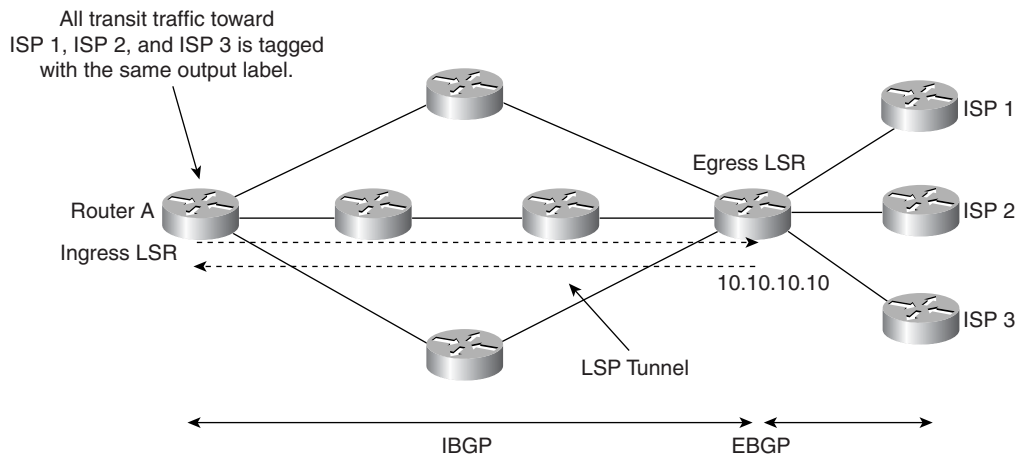
## RSVP LSP Tunnels

Service providers create LSP tunnels to aggregate traffic belonging to the same forwarding equivalency class. You have seen in Chapter 4, “Hybrid L2 and L3 in IP/MPLS Networks,” that multiple Virtual Private LAN Services (VPLSs) can be carried over a single LSP tunnel across the network.

LSPs are called LSP tunnels because the traffic going through an LSP tunnel is opaque to the intermediate LSRs between the ingress and egress LSRs. Figure 6-3 shows the establishment of an LSP tunnel between an ingress LSR and an egress LSR that is peering with multiple providers.

Notice how the LSP tunnel is formed using two unidirectional tunnels in both directions.

**Figure 6-3** LSP Tunnel Between Ingress and Egress LSRs



In Figure 6-3, traffic coming into router A and transiting the service provider’s network toward other service providers’ networks, such as ISP 1, ISP 2, and ISP 3, can all be grouped in the same forwarding equivalency class. This class is defined by all traffic destined for the exit router with IP address 10.10.10.10. In this case, all traffic toward 10.10.10.10 is tagged with the same outbound label at router A. This maps all transit traffic toward the same LSP tunnel.

The exit point for a given external route (10.10.10.10) is normally learned via the Internal Border Gateway Protocol (IBGP). After the traffic reaches the exit router, it is sent to the correct ISP, depending on the final external route.

## Label Binding and LSP Tunnel Establishment Via RSVP

RFC 3209, *RSVP-TE: Extensions to RSVP for LSP Tunnels*, defines the capabilities of extended RSVP. Regarding the operation of LSP tunnels, extended RSVP enables you to do the following:

- Perform downstream-on-demand label allocation, distribution, and binding.
- Observe the actual route traversed by an established LSP tunnel.
- Identify and diagnose LSP tunnels.
- Establish LSP tunnels with or without QoS requirements.
- Dynamically reroute an established LSP tunnel.
- Preempt an established LSP tunnel under administrative policy control.

To establish an LSP tunnel, the ingress LSR sends a PATH message to the egress LSR, which in turn replies with a reservation message (RESV). Upon completion of the handshake, an LSP tunnel is established. The PATH message indicates the PATH that the LSP should take, and the RESV message attempts to establish a bandwidth reservation following the opposite direction of the PATH message. PATH and RESV messages are explained in detail in the sections “Details of the PATH Message” and “Details of the RESV Message,” respectively.

RSVP-TE has defined new objects in support of creating LSP tunnels. These new objects, called `LSP_TUNNEL_IPv4` and `LSP_TUNNEL_IPv6`, help, among other things, identify LSP tunnels. The `SESSION` object, for instance, carries a tunnel ID, while the `SENDER_TEMPLATE` and `FILTER_SPEC` objects uniquely identify an LSP tunnel.

The following is the sequence of events needed to establish an LSP tunnel:

- 1 The first MPLS node on the path—that is, the ingress LSR (sender)—creates an RSVP PATH message with a session type of `LSP_TUNNEL_IPv4` or `LSP_TUNNEL_IPv6` and inserts a `LABEL_REQUEST` object into the PATH message.
- 2 The `LABEL_REQUEST` object indicates that a label binding for this path is requested and also indicates the network layer protocol that is to be carried over this path.

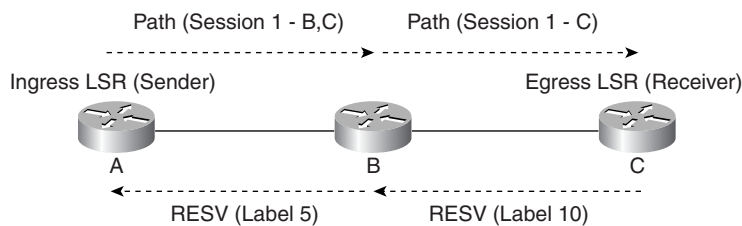
In addition to the `LABEL_REQUEST` object, the PATH message can carry a number of optional objects:

- **EXPLICIT\_ROUTE object (ERO)**—Specifies a predetermined path between the ingress and egress LSRs. When the ERO object is present, the PATH message is sent toward the first node indicated by the ERO, independent of the IGP shortest path.

- **RECORD\_ROUTE object (RRO)**—Used to record information about the actual route taken by the LSP. This information can be relayed back to the sender node. The sender node can also use this object to request notification from the network concerning changes in the routing path.
  - **SESSION\_ATTRIBUTE object**—Can be added to PATH messages to help in session identification and diagnostics. Additional control information, such as setup and hold priorities and local protection, is also included in this object.
- 3 The label allocation with RSVP is done using the downstream-on-demand label assignment mechanism.
  - 4 The RESV message is sent back upstream toward the sender, following the path created by the PATH message, in reverse order.
  - 5 Each node that receives an RESV message containing a LABEL object uses that label for outgoing traffic associated with this LSP tunnel.
  - 6 When the RESV message arrives at the ingress LSR, the LSP tunnel is established.

This process is illustrated in Figure 6-4.

**Figure 6-4** *Establishing an LSP Tunnel*



In Figure 6-4, ingress LSR A sends a PATH message toward LSR C with a session type object and an ERO. The ERO contains the explicit route that the PATH message needs to take. The ERO in this case is the set {B,C}, which dictates the path to be taken via LSR B, then LSR C.

In turn, LSR B propagates the PATH message toward LSR C according to the ERO. When LSR C receives the PATH message, it sends an RESV message that takes the reverse PATH indicated in the ERO toward LSR A. LSR C includes an inbound label of 10. Label 10 is used as an outbound label in LSR B. LSR B sends an RESV message toward LSR A with an inbound label of 5. Label 5 is used as an outbound label by LSR A. An LSP tunnel is formed between LSRs A and C. All traffic that is mapped to this LSP tunnel is tagged with label 5 at LSR A.

## Reservation Styles

The existing RSVP procedures for resource reservation can be applied on aggregated flows or traffic trunks. This model scales because it is done on trunks rather than flows and between



pairs of routers rather than pairs of hosts, as was originally intended for RSVP. The receiver node can select from among a set of possible reservation styles for each session, and each RSVP session must have a particular style. Senders have no influence on the choice of reservation style. The receiver can choose different reservation styles for different LSPs. Bandwidth reservation is not mandatory for the operation of RSVP-TE. It is up to the service provider to engineer the networks as necessary to meet the SLAs.

The following sections discuss the different reservation styles listed here and their advantages and disadvantages:

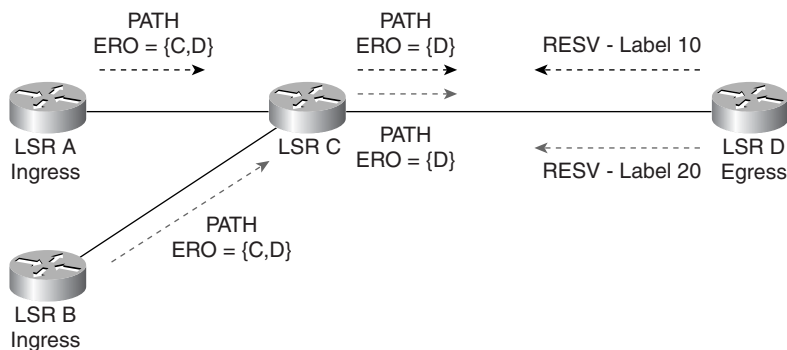
- Fixed Filter (FF)
- Shared Explicit (SE)
- Wildcard Filter (WF)

### Fixed Filter Reservation Style

The FF reservation style creates a distinct reservation for traffic from each sender. This style is normally used for applications whose traffic from each sender is independent of other senders. The total amount of reserved bandwidth on a link for sessions using FF is the sum of the reservations for the individual senders. Because each sender has its own reservation, a unique label is assigned to each sender. This can result in a point-to-point LSP between every sender/receiver pair. An example of such an application is a one-on-one videoconferencing session. Bandwidth reservations between different pairs of senders and receivers are independent of each other.

In Figure 6-5, ingress LSRs A and B create distinct FF-style reservations toward LSR D. The total amount of bandwidth reserved on link C-D is equal to the sum of reservations requested by A and B. Notice also that LSR D has assigned different labels in the RESV messages toward A and B. Label 10 is assigned for sender A, and label 20 is assigned for sender B. This creates two distinct point-to-point LSPs—one between A and D and the other between B and D.

**Figure 6-5** Fixed Filter Reservation Style



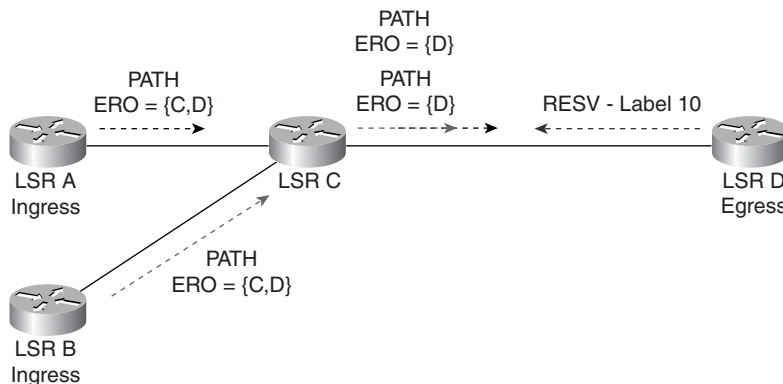
## Shared Explicit Reservation Style

The SE reservation style allows a receiver to explicitly select a reservation for a group of senders—rather than one reservation per sender, as in the FF style. Only a single reservation is shared between all senders listed in the particular group.

SE style reservations can be provided using one or more multipoint-to-point LSPs per sender. Multipoint-to-point LSPs may be used when PATH messages do not carry the ERO, or when PATH messages have identical EROs. In either of these cases, a common label may be assigned.

PATH messages from different senders can each carry their own ERO, and the paths taken by the senders can converge and diverge at any point in the network topology. When PATH messages have differing EROs, separate LSPs for each ERO must be established. Figure 6-6 explains the SE style even further.

**Figure 6-6** *Shared Explicit Reservation Style*



In Figure 6-6, LSRs A and B are using the SE style to establish a session with LSR D. The reservation for link C-D is shared between A and B. In this example, both PATH messages coming from A and B have the same ERO and are converging on node C. Notice that D has allocated a single label 10 in its RESV message, hence creating the multipoint-to-point LSP. An example of such an application is a videoconferencing session between multiple branch offices in Europe and the main office in the United States. The bandwidth reserved on the international link is set for a certain amount, and the number of remote branch offices is set in a way that the total amount of bandwidth used by the branch offices does not exceed the total reserved bandwidth.

## Wildcard Filter Reservation Style

A third reservation style that is defined by RSVP is the WF reservation style. Unlike the SE style, where the receiver indicates the specific list of senders that are to share a reservation, with the WF reservation style, a single shared reservation is used for all senders to a session. The total reservation on a link remains the same regardless of the number of senders.

This style is useful for applications in which not all senders send traffic at the same time. If, however, all senders send simultaneously, there is no means of getting the proper reservations made. This restricts the applicability of WF for TE purposes.

Furthermore, because of the merging rules of WF, EROs cannot be used with WF reservations. This is another reason that prevents the use of the WF style for traffic engineering.

## Details of the PATH Message

The PATH message can include several different RSVP objects, including the following:

- LABEL\_REQUEST
- EXPLICIT\_ROUTE
- RECORD\_ROUTE
- SESSION\_ATTRIBUTE
- FLOW\_SPEC
- SENDER\_TEMPLATE
- SESSION

Figure 6-7 shows the format of the PATH message.

**Figure 6-7** *RSVP-TE PATH Message*

Common Headers
SESSION Object
EXPLICIT_ROUTE Object (ERO)
LABEL_REQUEST Object
RECORD_ROUTE Object (RRO)
SESSION_ATTRIBUTE Object
SENDER_TEMPLATE Object
FLOW_SPEC Object

The following sections describe each object in more detail.

### LABEL\_REQUEST Object

The LABEL\_REQUEST object is used to establish label binding for a certain path. It also indicates the network layer protocol that is to be carried over this path. The reason for this is

that the network layer protocol sent down an LSP does not necessarily have to be IP and cannot be deduced from the L2 header, which only identifies the higher-layer protocol as MPLS. The LABEL\_REQUEST object has three possible C\_Types (Class\_Types):

- **Type 1, label request without label range**—This is a request for a regular 32-bit MPLS label that sits in the shim layer between the data link and network layer headers.
- **Type 2, label request with an ATM label range**—This request specifies the minimum and maximum virtual path identifier (VPI) and virtual connection identifier (VCI) values that are supported on the originating switch. This is used when the MPLS label is carried in an ATM header.
- **Type 3, label request with Frame Relay label range**—This request specifies the minimum and maximum data-link connection identifier (DLCI) values that are supported on the originating switch. This is used when the MPLS label is carried in a Frame Relay header.

When the PATH message reaches an LSR, the LABEL\_REQUEST object gets stored in the path state block for further use by refresh messages. When the PATH message reaches the receiver, the presence of a LABEL\_REQUEST object triggers the receiver to allocate a label and to place the label in the LABEL object for the corresponding RESV message. If a label range is specified, the label must be allocated from that range. Error messages might occur in cases where the receiver cannot assign a label, cannot recognize the protocol ID, or cannot recognize the LABEL\_REQUEST object.

## EXPLICIT\_ROUTE Object

The EXPLICIT\_ROUTE object (ERO) is used to specify an explicit path across the network independent of the path specified by the IGP. The contents of an ERO are a series of variable-length data items called *subobjects*.

A subobject is an abstract node that can be either a single node or a group of nodes such as an autonomous system. This means that the explicit path can cross multiple autonomous systems, and the hops within each autonomous system are opaque (hidden) from the ingress LSR for that path.

The subobject contains a 1-bit Loose Route field (L). If set to 1, this field indicates that the subobject is a loose hop in the explicit path, and if set to 0, it indicates that the subobject is a strict hop. A strict hop indicates that this hop is physically adjacent to the previous node in the path.

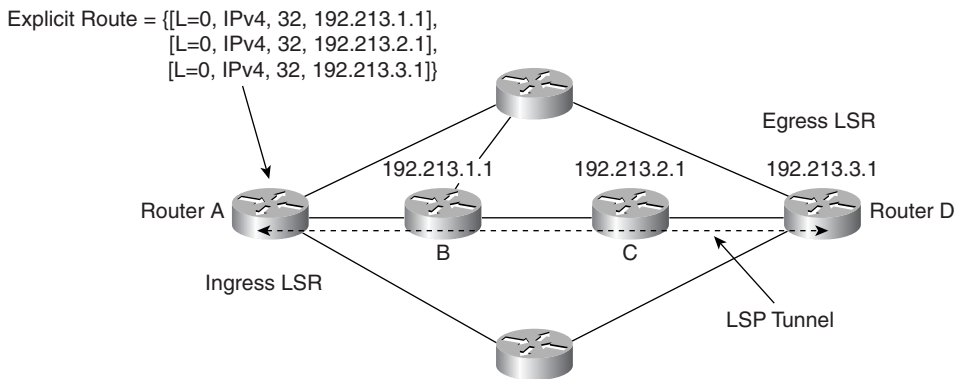
The subobject also contains a Type field, which indicates the types of the content subobjects. Some defined values of the Type field are as follows:

- **1: IPv4 Prefix**—Identifies an abstract node with a set of IP prefixes that lie within this IPv4 prefix. A prefix of length 32 is a single node (for example, a router's IP loopback address).
- **2: IPv6 Prefix**—Identifies an abstract node with a set of IP prefixes that lie within this IPv6 prefix. A prefix of length 128 is a single node (for example, a router's IP loopback address).
- **32: Autonomous System number**—Identifies an abstract node consisting of the set of nodes belonging to the autonomous system.

Figures 6-8 and 6-9 illustrate two scenarios in which an explicit path is being established using strict and loose subobjects, respectively, of the Type IPv4 prefix and with a subobject length of 32.

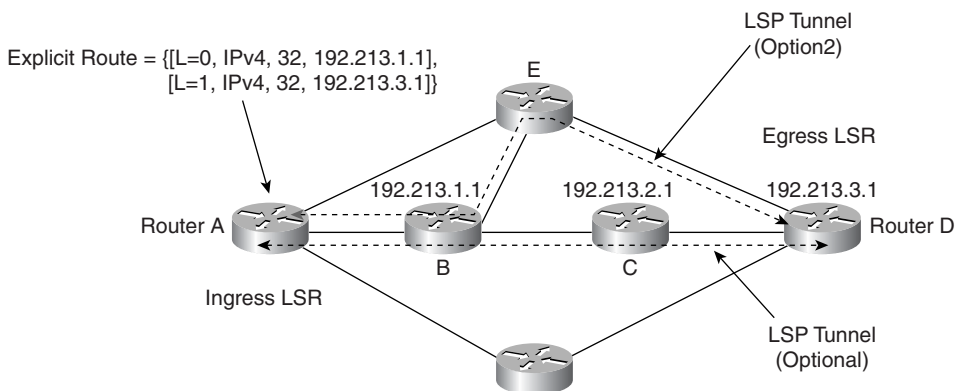
In Figure 6-8, ingress LSR A sends a PATH message toward LSR D with an ERO that indicates a strict hop across routers B (192.213.1.1), C (192.213.2.1), and D (192.213.3.1). When B receives the PATH message, it propagates it toward C, and C propagates the message toward D. In turn, D sends a RESV message to A along the same path, and the label binding takes place. The ERO itself is modified at each hop. Each node in the ERO list removes itself from the ERO as the PATH message is forwarded.

**Figure 6-8** *Explicit Route, Strict Hops*



In Figure 6-9, ingress LSR A sends a PATH message toward LSR D with an ERO that indicates a strict hop toward B. From router B, a loose hop is used. When router B receives the PATH message, it would send the PATH message to D along any available route. In this example, there are two possible routes toward D—one via a direct connection to C and the other via router E. The way the loose hop is picked depends on the IGP route that is available toward D.

**Figure 6-9** *Explicit Route, Loose Hops*



It is important to note that intermediate LSRs between the sender and receiver may also change the ERO by inserting subobjects. An example is where an intermediate node replaces a loose route subobject with a strict route subobject to force the traffic around a specific path. Also, the presence of loose nodes in an explicit route implies that it is possible to create forwarding loops in the underlying routing protocol during transients. Loops in an LSP tunnel can be detected using the RECORD\_ROUTE object (RRO), as discussed in the next section.

## RECORD\_ROUTE Object

The RRO is used to collect detailed path information and is useful for loop detection and for diagnostics. By adding an RRO to the PATH message, the sender can receive information about the actual PATH taken by the LSP. Remember that although the ERO specifies an explicit PATH, the PATH might contain loose hops, and some intermediate nodes might change the ERO, so the final PATH recorded by the RRO could be different from the ERO specified by the sender. The RRO can be present in both RSVP PATH and RESV messages. The RRO is present in an RESV message if the RRO that has been recorded on the PATH message needs to be returned to the ingress LSR.

There are three possible uses of RROs in RSVP:

- **Loop detection**—An RRO can function as a loop-detection mechanism to discover L3 routing loops or loops inherent in the explicit route.
- **Path information collection**—An RRO collects up-to-date detailed path information hop-by-hop about RSVP sessions, providing valuable information to the sender or receiver. Any path change (because of network topology changes) is reported.
- **Feedback into ERO**—An RRO can be used as input to the ERO object. If the sender receives an RRO via the RESV message, it can alter its ERO in the next PATH message. This can be used to “pin down” a session path to prevent the path from being altered even if a better path becomes available.

The initial RRO contains only one subobject: the sender’s IP addresses. When a PATH message containing an RRO is received by an intermediate router, the router stores a copy of it in the path state block. When the PATH message is forwarded to the next hop, the router adds to the RRO a new subobject that contains its own IP address. When the receiver sends the RRO to the sender via the RESV message, the RRO has the complete route of the LSP from ingress to egress.

## SESSION\_ATTRIBUTE Object

The SESSION\_ATTRIBUTE object allows RSVP-TE to set different LSP priorities, preemption, and fast-reroute features. These are used to select alternate LSPs in case of a failure in the network. The SESSION\_ATTRIBUTE is carried in the PATH message. It includes fields such as Setup Priority and Holding Priority, which affect whether this session can preempt or can be preempted by other sessions. A Flag field is also used to introduce options such as whether transit routers can use local mechanisms that would violate the ERO and cause local

repair. Other Flag options indicate that the tunnel ingress node may choose to reroute this tunnel without tearing it down.

## FLOW\_SPEC Object

An elementary RSVP reservation request consists of a FLOW\_SPEC together with a FILTER\_SPEC; this pair is called a *flow descriptor*. The FLOW\_SPEC object specifies a desired QoS. The FILTER\_SPEC object, together with a SESSION object specification, defines the set of data packets—the “flow”—to receive the QoS defined by the flowspec. An example of the use of FLOW\_SPEC with RSVP-TE would be to indicate which path certain traffic gets put on based on the QoS characteristics of such traffic. Data packets that are addressed to a particular session but that do not match any of the filter specs for that session are handled as best-effort traffic. The flowspec in a reservation request generally includes a service class and two sets of numeric parameters:

- An Rspec (R for “reserve”) that defines the desired QoS
- A Tspec (T for “traffic”) that describes the data flow

## SENDER\_TEMPLATE Object

PATH messages are required to carry a SENDER\_TEMPLATE object, which describes the format of data packets that this specific sender originates. This template is in the form of a FILTER\_SPEC that is typically used to select this sender’s packets from others in the same session on the same link. The extensions of RSVP for TE define a new SENDER\_TEMPLATE C-Type (LSP\_TUNNEL\_IPv4) that contains the IPv4 address for the sender node and a unique 16-bit identifier, the LSP\_ID, that can be changed to allow a sender to share resources with itself. This LSP\_ID is used when an LSP tunnel that was established with an SE reservation style is rerouted.

## SESSION Object

The SESSION object is added to the PATH message to help identify and diagnose the session. The new LSP\_TUNNEL\_IPv4 C-Type contains the IPv4 address of the tunnel’s egress node and a unique 16-bit identifier that remains constant over the life of the LSP tunnel, even if the tunnel is rerouted.

## Details of the RESV Message

An RESV message is transmitted from the egress LSR toward the ingress in response to the receipt of a PATH message. The RESV message is used for multiple functions, including: distributing label bindings, requesting resource reservations along the path, and specifying the reservation style (FF or SE).

The RSVP RESV message can contain a number of different objects such as LABEL, RECORD\_ROUTE, SESSION, and STYLE.

Figure 6-10 shows the format of the RESV message.

**Figure 6-10** *RSVP-TE RESV Message*

Common Headers
SESSION Object
LABEL Object
RECORD_ROUTE Object (RRO)
STYLE Object (FF or SE)
<Filter Descriptor Lists>

The RECORD\_ROUTE and SESSION objects were described as part of the PATH message in the preceding section. The LABEL object contains the label or stack of labels that is sent from the downstream node to the upstream node. The STYLE object specifies the reservation style used. As you have learned, the FF and SE reservation styles filters are used for TE. For the FF and SE styles, a label is provided for each sender to the LSP.

## Understanding MPLS Fast Reroute

One of the requirements for TE is the capability to reroute an established TE tunnel under various conditions. Such rerouting capabilities could include the following:

- Setting administrative policies to allow the LSP to reroute, such as when the LSP does not meet QoS requirements.
- Rerouting an LSP upon failure of a resource along the TE tunnel's established path.
- Setting an administrative policy that might require that an LSP that has been rerouted must return to its original path when a failed link or router becomes available.

Network operation must not be disrupted while TE rerouting is in progress. This means that you need to establish backup tunnels ahead of time and transfer traffic from the old tunnel to the new tunnel before you tear down the old tunnel. This concept is called *make-before-break*. A problem could arise if the old and new tunnels are competing for network resources; this might prevent the new tunnel from being established, because the old tunnel that needs to be torn down still has the allocated resources.

One of the advantages of using RSVP-TE is that the protocol has many hooks to take care of such problems. RSVP uses the SE reservation style to prevent the resources used by an old tunnel from being released until the new tunnel is established. The SE reservation style also prevents double counting of the resources when moving from an old tunnel to a new tunnel.

The speed of rerouting a failed tunnel is crucial for maintaining SLAs for real-time applications in the metro. When an LSP tunnel fails, the propagation of the failure to the ingress LSR/LER that established the tunnel and the convergence of the network to a new LSP tunnel could cause higher-level applications to time out. MPLS fast reroute allows an LSP tunnel to be rerouted in tens of milliseconds.



RSVP-TE can be used to establish backup LSP tunnels if active LSP tunnels fail. There are two methods of doing so:

- End-to-end repair
- Local repair

## End-to-End Repair

With the end-to-end repair method, the whole LSP tunnel is backed up from the ingress LSR to the egress LSR. If the LSP fails because of a break in the network, a whole new LSP is established end to end. In this case, it is possible to presignal the secondary path, which is quicker than resignaling the LSP.

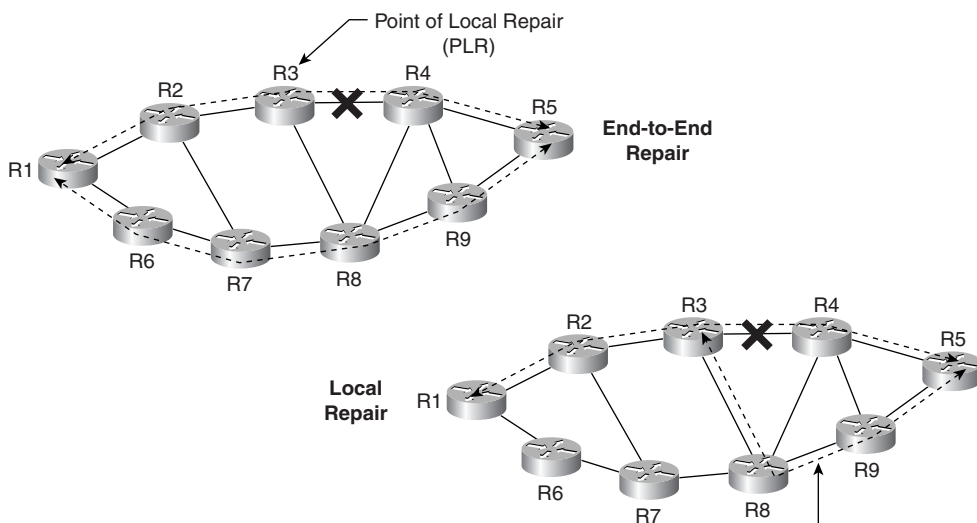
## Local Repair

Local repair allows the LSP to be repaired at the place of failure. This allows the existing LSP to reroute around a local point of failure rather than establish a new end-to-end LSP. The benefit of repairing an LSP at the point of failure is that it decreases the network convergence time and allows the traffic to be restored in tens of milliseconds. This is important to meet the needs of real-time applications such as Voice over IP or video over IP, which are the next-generation services for metro networks.

To achieve restoration in tens of milliseconds, backup LSPs are signaled and established in advance of failure. The traffic is also redirected as close to the failure as possible. This reduces the delays caused by propagating failure notification between LSRs.

Figure 6-11 shows the difference between using local repair and end-to-end repair.

**Figure 6-11** *The Value of Local Repair*



JUNIPER Exhibit 1003  
App. 6, pg. 166

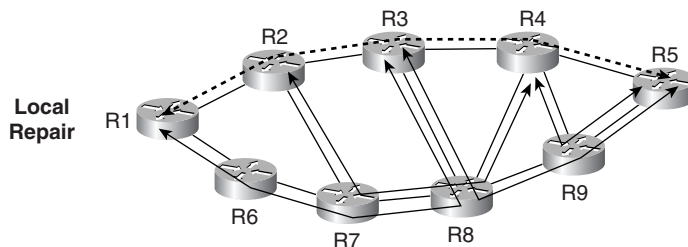
In Figure 6-11, an LSP tunnel is established between R1 and R5. If end-to-end repair is used and a failure occurs anywhere on the links or routers between R1 and R5—the R3-R4 link in this example—failure notification has to propagate from R3 all the way to R1. Also, all the LSRs, including R1 and R2, have to be involved in recomputing the new path. If the secondary path is pre signaled between R1 and R5, convergence occurs much faster.

Conversely, local repair allows the traffic to be redirected closest to the failure and hence dramatically reduces the restoration time. If local repair is used, the LSP could be spliced between R3 and R5, bypassing the failure. Of course, this is all great as long as you know where the failure will occur so that you can work around it. Because this is impossible to know, you have to predict which links are carrying critical data and need to be protected. Two local repair techniques, one-to-one backup and facility backup, are discussed next.

### One-to-One Backup

In the one-to-one backup method, a node is protected against a failure on its downstream link or node by creating an LSP that starts upstream of that node and intersects with the original LSP somewhere downstream of the point of link or node failure. In Figure 6-11 (local repair) the one-to-one backup method was used to protect against a failure of the link R3-R4, or the failure of node R4. In this case, R3's backup is an LSP that starts at R3 and ends downstream of the R3-R4 link on the R5 node. The partial LSP that starts from R3 and goes around R4 and splices back into the original LSP is called a *detour LSP*. To fully protect an LSP that passes N nodes, there could be as many as N-1 detours. In the example in Figure 6-12, to protect the LSP between R1 and R5, there could be as many as four detour LSPs.

**Figure 6-12** Full LSP Protection



The LSP that needs to be protected is R1-R2-R3-R4-R5:

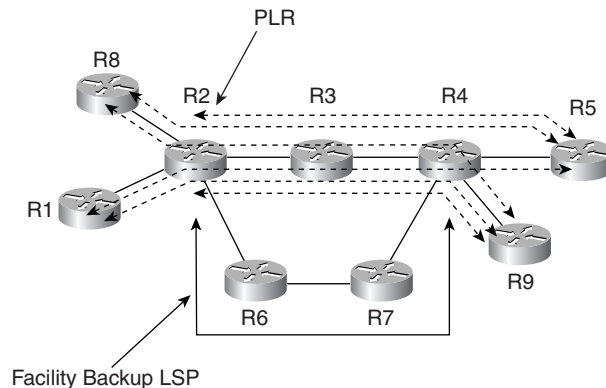
- Upon failure of the R1-R2 link, or R2 node, R1's detour LSP would be R1-R6-R7-R8-R3.
- Upon failure of the R2-R3 link, or R3 node, R2's detour LSP would be R2-R7-R8-R4.
- Upon failure of the R3-R4 link, or R4 node, R3's detour LSP would be R3-R8-R9-R5.
- Upon failure of the R4-R5 link, R4's detour LSP would be R4-R9-R5.

The point (router) that initiates the detour LSP is called the *point of local repair (PLR)*. When a failure occurs along the protected LSP, the PLR redirects the traffic onto the local detour. If R1-R2 fails, R1 switches the traffic into the detour LSP R1-R6-R7-R8-R3.

## Facility Backup—Bypass

Another method for protecting the LSP against failure is called the *facility backup*. Instead of creating a separate LSP for every backed-up LSP, a single LSP is created that serves to back up a set of LSPs. This LSP is called a *bypass tunnel*. The bypass tunnel intersects the path of the original LSPs downstream of the PLR. This is shown in Figure 6-13.

**Figure 6-13** *Bypass Tunnel*




The bypass tunnel R2-R6-R7-R4 is established between R2 and R4. The scalability improvement from this technique comes from the fact that this bypass tunnel can protect any LSP from R1, R2, and R8 to R4, R5, and R9. As with the one-to-one technique, to fully protect an LSP that traverses  $N$  nodes, there could be as many as  $N-1$  bypass tunnels. However, each of these bypass tunnels can protect a set of LSPs.

## Conclusion

This chapter has discussed the basics of RSVP-TE and how it can be applied to establish LSPs, bandwidth allocation, and fast-reroute techniques. A detailed explanation of the RSVP-TE messages and objects was offered to give you a better feel for this complex protocol, which probably requires a book of its own. Many of the techniques explained in this chapter apply to provisioning scalable L2 metro Ethernet services.

The metro will consist of a mix of technologies ranging from Ethernet switches to SONET/SDH equipment to optical switches. Creating a unified control plane that is capable of provisioning LSP tunnels end to end and helping in the configuration and management of such equipment becomes crucial. You have seen the MPLS control plane used for packet networks. The flexibility and standardization of MPLS is extending its use to TDM and optical networks. The next two chapters discuss Generalized MPLS (GMPLS) and how this control plane becomes universal in adapting not only to packet networks but also across TDM and optical networks.



This chapter covers the following topics:

- Understanding GMPLS
- Establishing the Need for GMPLS
- Signaling Models
- Label Switching in a Nonpacket World

# MPLS Controlling Optical Switches

---

The operation of today's optical networks is manual and operator-driven, which increases network operational complexities and cost. The industry has been looking for methods that reduce the operational burden of manual circuit provisioning, reduce costs, and offer a more dynamic response to customer requirements. In other words, the industry wants to be able to deploy time-division multiplexing (TDM) and optical circuits more dynamically and wants faster provisioning times.

The principles upon which MPLS technology is based are generic and applicable to multiple layers of the transport network. As such, MPLS-based control of other network layers, such as the TDM and optical layers, is also possible. The Common Control and Measurement Plane (CCAMP) Working Group of the IETF is currently working on extending MPLS protocols to support multiple network layers and new TDM and optical services. This concept, which was originally referred to as Multiprotocol Lambda Switching (MP $\lambda$ S), is now referred to as Generalized MPLS (GMPLS). This chapter refers to definitions from the CCAMP Working Group in the areas that cover the GMPLS architecture and concepts.

## Understanding GMPLS

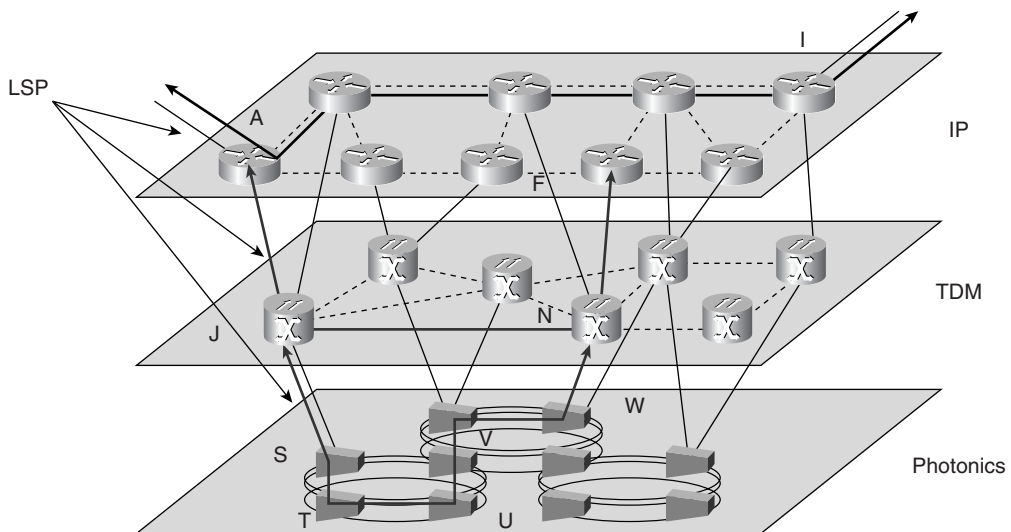
Generalized MPLS is a set of architectures and protocols that enables TDM and optical networks to behave more dynamically. GMPLS builds on the MPLS control, which is well known and proven to work, and extends the capabilities of MPLS to control TDM and optical networks, including TDM switches, wavelength switches, and physical port switches.

In the same way that MPLS builds label switched paths (LSPs) between packet switches, GMPLS extends the concept of LSPs to TDM and optical switches. Figure 7-1 illustrates a three-layer hierarchy where GMPLS LSPs are built between two points in the network over multiple layers.

Figure 7-1 shows how the MPLS LSP concept that is used for the IP packet/cell layer can be extended to address the TDM and optical layers. On the IP layer, an LSP is formed between routers A and I. On the TDM layer, an LSP is formed between SONET/SDH multiplexers J and N. On the photonics layer, an LSP is formed between optical switches

S and W along the path S-T-U-V-W. The establishment of LSPs of course necessitates that TDM and optical switches become aware of the GMPLS control plane while still using their own multiplexing and switching techniques. This is one of the powerful advantages of MPLS, because the control and forwarding planes are decoupled.

Figure 7-1 GMPLS LSPs



GMPLS has two applications, both of which can be used in metro network deployments. First, for dynamic circuit provisioning, GMPLS can be used to establish point-to-point or multipoint-to-point virtual private optical networks. Second, GMPLS can be used for protection on the circuit level. In the context of deploying Ethernet services over an optical cloud, GMPLS would extend across L2 Ethernet switches/routers, SONET/SDH multiplexers, and optical cross-connects (OXC) to establish end-to-end circuits. Note that such deployments have not occurred yet, and it is unclear at the moment how fast or slow the adoption of GMPLS will evolve. The next section describes in more detail the need for GMPLS in optical networks.

## Establishing the Need for GMPLS

Anyone who has been in the networking industry for a while would likely raise the issue of whether GMPLS is really needed or is overkill. After all, we have managed so far to build large-scale TDM networks with all sorts of methods, and we have seen improvement in tools to facilitate the operation and management of those networks. To understand the issue of whether GMPLS is necessary, you need to look first at how TDM networks function today and then at how they could benefit from GMPLS.

The following section describes the provisioning model of today's network deployments, which are more static with centralized management. The problem with this model is that it doesn't enable carriers to provide new services that involve the dynamic establishment and restoration of TDM and optical circuits while minimizing the operational cost and provisioning times. This is the problem that the GMPLS model attempts to address. If GMPLS could solve this problem, the result would be a better service experience for customers and increased revenue for the carrier. However, adopting GMPLS would also require fundamental changes to the way you administer, manage, and build networks.

## Static and Centralized Provisioning in TDM Networks

Currently, TDM and optical networks are statically provisioned. Provisioning a point-to-point circuit takes weeks to accomplish, because it entails lengthy administrative and architectural tasks. The majority of today's TDM network management and provisioning models are centralized. Provisioning is done either manually or with automated tools and procedures that reside in a central network management entity that has knowledge of the whole network and its elements. To handle scalability issues, such as having too many nodes (thousands) to manage, network managers use a hierarchical approach in which they manage multiple domains separately and higher management layers oversee the whole service operation. The network topology includes topology information about rings and meshed networks. The network resources include information about the network elements, such as fibers, ducts, links, and their available capacity. Entering such information manually is tedious and error-prone, especially in networks that require constant changes for expansion and upgrades.

The provisioning process involves the following:

- **Administrative tasks**—Request for a circuit involves the paperwork or web-based tools for a customer such as a large enterprise to fill out and submit as a request for a circuit. The request is fulfilled by the network operator.
- **Network planning**—The network operator has to run simulations to find out whether the network has the capacity to absorb the additional circuits and to determine how to optimize the network resources. This task is normally done on a set of circuits at regular intervals. Network planning touches different parts of the network, depending on where these circuits start and end. High-capacity circuits normally put a major strain on metropolitan area networks that were built for traditional voice services. In some cases, the addition of one TDM circuit might cause the operator to build more metro SONET rings to absorb additional capacity; hence, the service could be rejected or delayed until the operator justifies the economics of building more circuits for a particular customer.
- **Installing the physical ports**—This is the manual task of installing the WAN ports at the customer premises and installing the connection/circuit between the customer premises and the operator's networks.

- **Circuit provisioning**—This is the task of establishing the circuit end to end, using either management tools or manual configuration. Circuit provisioning is one of the most challenging areas because it requires establishing circuits across multiple components, sometimes from different vendors, with different interfaces and different protocols. Circuit provisioning also involves testing the circuit to see whether it complies with the SLA that was promised to the customer.
- **Billing**—As simple as it may sound, a service cannot be deployed until it can be billed for. Whether flat billing or usage-based billing is used, the task of defining and accounting for the right variables is not simple.
- **Network management**—Last but not least is the continuous process of managing the different network elements, keeping the circuits up and running, and restoring the circuits in case of network failures.

## The Effect of a Dynamic Provisioning Model

GMPLS offers a dynamic provisioning model for building optical networks. In this more dynamic and decentralized model, information about the network topology and resources can be exchanged via protocols such as OSPF traffic engineering (OSPF-TE) and IS-IS traffic engineering (ISIS-TE). The information is available to all nodes in the network, including the Network Management System (NMS), which can act upon it. The dissemination of such information via the routing protocols gives the operator a clearer view of the network, which facilitates planning, provisioning, and operation.

Figures 7-2 and 7-3 show two scenarios of centralization and decentralization.

**Figure 7-2** *Centralized, Static Control*

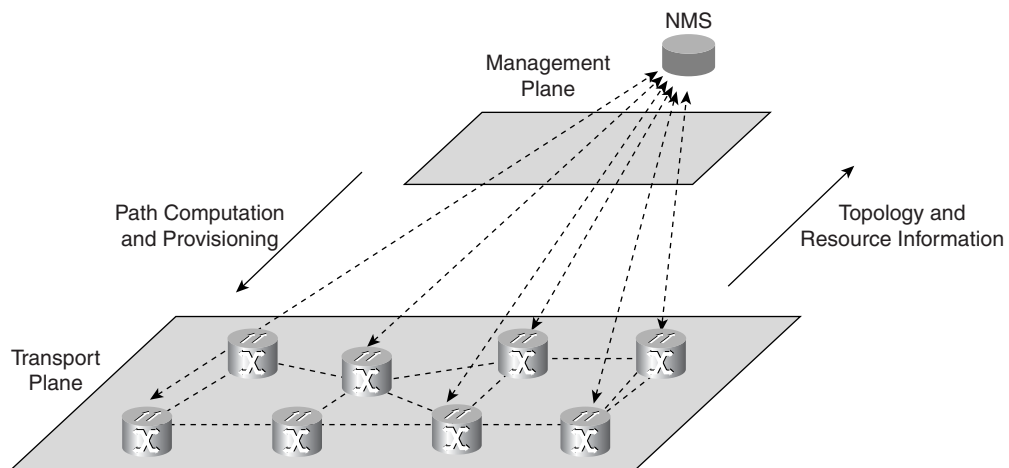




Figure 7-2 shows the centralized approach, in which all nodes communicate with the NMS and relay information about topology and resources to a central database. The NMS acts on this information for path computation and provisioning.

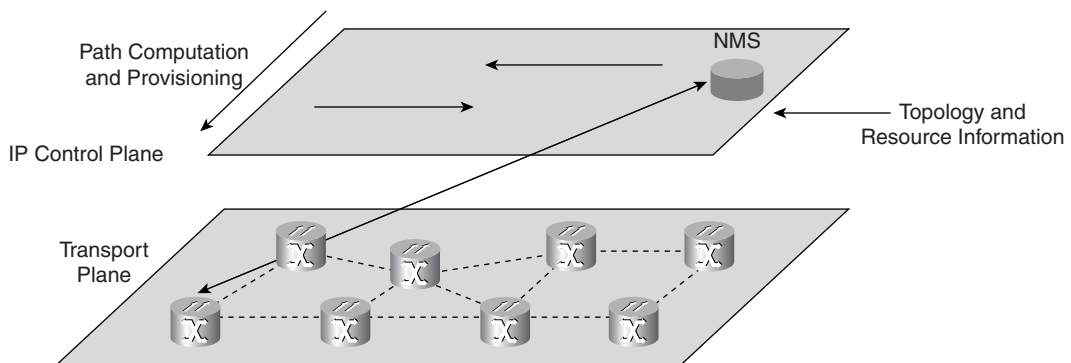
In optical networks, the control plane can be exchanged between the different network systems (optical switches, routers, and so on) via in-band or out-of-band communications. The GMPLS control plane can use multiple communication models:

- Over a separate fiber
- Over a separate wavelength
- Over an Ethernet link
- Over an IP tunnel through a separate management network
- Over the overhead bytes of the data-bearing link

A communication over a SONET/SDH data communication channel (DCC), for example, could use the SONET/SDH DCC path D1-D3 or the line D4-D24 overhead bytes. For wavelength-division multiplexing (WDM) nodes, a separate wavelength could be dedicated as an IP management channel. It is important that the management channel be operational at all times. If, for example, the management is done in-band, a network failure could cause the management channel to fail. Hence, the nodes and links could become inaccessible and couldn't be restored.

Figure 7-3 shows the decentralized approach, in which the nodes exchange topology and resource information via different protocols (for example, OSPF-TE) through the IP control plane running in-band or out-of-band. Path computation and provisioning can be triggered dynamically or via an NMS station. The NMS station could simply send commands to one of the ingress nodes to initiate a path.

**Figure 7-3** *Decentralized, Dynamic Control*



When applying routing to circuit-switched networks, it is useful to compare and contrast this situation with the IP packet routing case, which includes the following two scenarios:

- Topology and resource discovery
- Path computation and provisioning

## Topology and Resource Discovery

In the case of routing IP packets, all routes on all nodes must be calculated exactly the same way to avoid loops and “black holes.” Conversely, in circuit switching, routes are established per circuit and are fixed for that circuit. To accommodate the optical layer, routing protocols need to be supplemented with new information, such as available link capacity. Due to the increase in information transferred in the routing protocol, it is important to separate a link’s relatively static parameters from those that may be subject to frequent changes.

Using a dynamic model to report link capacity in TDM and optical networks can be challenging. You have to find a balance where you are getting accurate reports about specific signals without flooding the network with too much information.

## Path Computation and Provisioning

In packet networks, path computation and reachability are very dynamic processes. Routing protocols determine the best path to a destination based on simple metrics such as link bandwidth. As described in Chapter 6, “RSVP for Traffic Engineering and Fast Reroute,” MPLS with RSVP-TE gives you more control to traffic-engineer the network. For optical networks, path computation and provisioning depend on the following information:

- The available capacity of the network links
- The switching and termination capabilities of the nodes and interfaces
- The link’s protection properties

When such information is exchanged dynamically via routing protocols, the network always has a real-time view of link and node capacity and properties that can be used to calculate the most suitable path.

With all the required tasks for deploying a service, optimizing the right mix of tasks becomes challenging. No one solution has a positive impact on all variables at the same time. Applying a dynamic provisioning model to the network, for example, would shorten provisioning but would also make network planning, service billing, and network management more challenging. After all, carriers have always dealt with a static provisioning and TE model, because they have always had total control of the network, its resources, and its behavior. Besides, for legacy SONET equipment that does not have the capability to run GMPLS and dynamic protocols, static approaches remain necessary. As such, a combination of static and dynamic, centralized and decentralized approaches would apply to most network designs.

The transition to adopting GMPLS will take many steps and will happen faster with some providers than others. Adopting this new model will be much easier for alternative providers and greenfield operators than for incumbents, which have well-defined procedures and tools that have been used for years. The cost justification for adopting GMPLS is not yet as clear as its benefits are. The next section discusses the dynamic provisioning model in more detail.

To adapt MPLS to control TDM and optical networks, the following primary issues need to be addressed:

- Addressing
- Signaling
- Routing
- Restoration and survivability

The following section begins by looking at the different signaling models that are in use and that are proposed for optical networks. Chapter 8, “GMPLS Architecture,” provides more details about the rest of the topics in the preceding list.

## Signaling Models

Signaling is a critical element in the control plane. It is responsible for establishing paths along packet-switched capable (PSC) and non-PSC networking devices such as routers, TDM cross-connects, and OXCs. PSC networks have no separation between the data and signaling paths; both data traffic and control traffic are carried over the same channels. In optical networks, control traffic needs to be separated from data traffic. One of the reasons is that OXCs are transparent to the data, because they perform light or lambda switching, whereas control traffic needs to be terminated at each intermediary OXC, because it carries the information to manage the data flows and information exchange between OXCs.

Multiple proposals exist for a signaling infrastructure over optical networks. The most common models are the following:

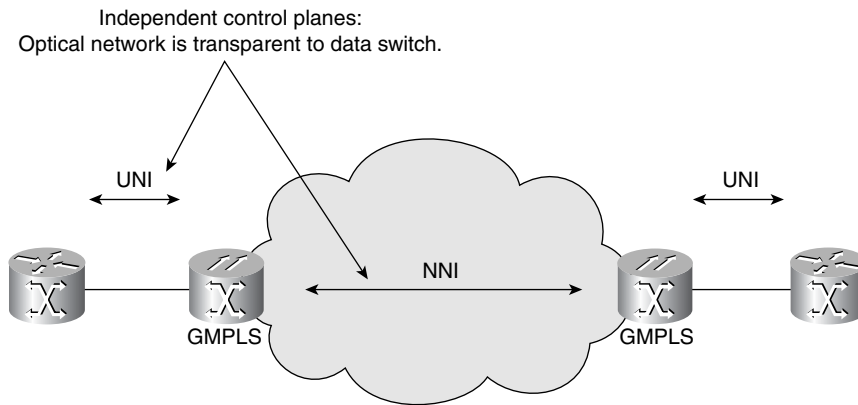
- The overlay model
- The peer model
- The augmented model

### The Overlay Model

In this model, illustrated in Figure 7-4, the internals of the optical infrastructure are totally transparent to the data-switching infrastructure. The optical infrastructure is treated as a separate intelligent network layer. Data switches at the edges of the optical infrastructure can statically or dynamically provision a path across the optical cloud. This is very similar to the IP-over-ATM model that exists today in carrier backbones. In this model, two independent control planes exist:

- **Within the packet layer**—The control plane runs on the User-to-Network Interface (UNI) between the data switches at the edge of the optical cloud and the optical switches.
- **Within the optical network**—The control plane runs on the Network-to-Network Interface (NNI) between the optical switches.

**Figure 7-4** *Overlay Model*

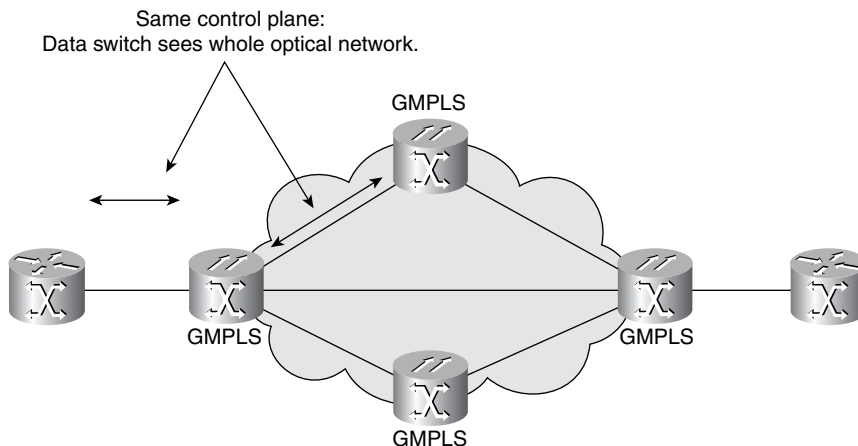


The overlay model applies in environments with limited or unknown trust that apply strict levels of policy and authentication and that limit routing information transfer.

## The Peer Model

In the peer model, illustrated in Figure 7-5, the IP/MPLS layers act as peers of the optical transport network, such that a single control plane runs over both the IP/MPLS and optical domains. As far as routing protocols are concerned, each edge device is adjacent to the optical switch it is attached to. The label switch routers (LSRs) and OXC's exchange complete information. The routers/data switches know the full optical network topology and can compute paths over it. For data-forwarding purposes, a full optical mesh between edge devices is still needed so that any edge node can communicate with any other edge node.

**Figure 7-5** *Peer Model*

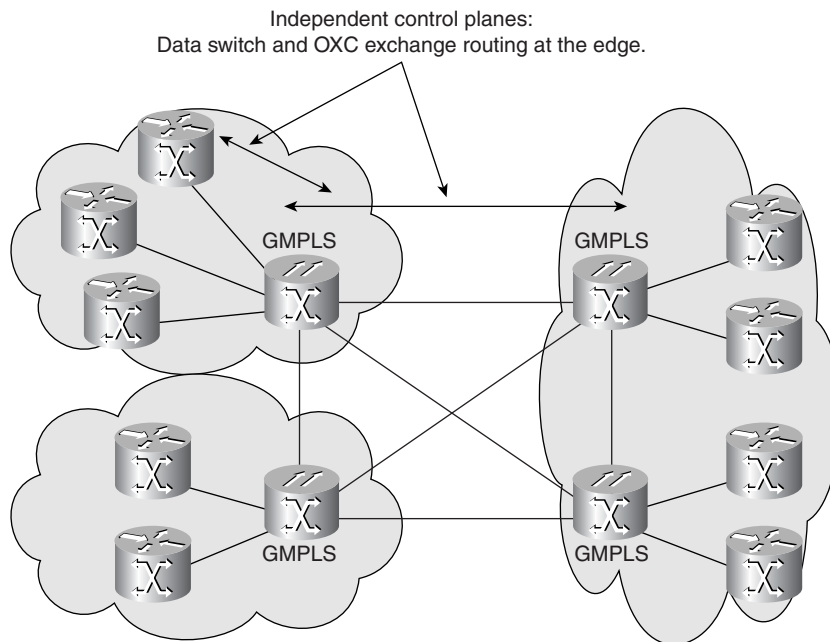


The advantage of the peer model is that, by developing uniform control, it gives the IP layer visibility into the optical layer and supports better IGP scaling if routers are meshed over an operational network. The peer model is much more similar to the use of MPLS than is an IP-over-ATM overlay model.

## The Augmented Model

The augmented model, illustrated in Figure 7-6, is a hybrid model that falls between the overlay and peer models. In the augmented model, separate control planes for the optical and IP domains are used, but some edge data switches still could have a limited exchange of routing information with border optical switches. This model allows for a transition from the overlay model to the more evolved peer model. One possible scenario in which the augmented model could be used is where a provider owns the data switches and the border optical switches and relies on a transport service offered by a different provider that owns the core optical switches.

**Figure 7-6** *Augmented Model*



## Label Switching in a Nonpacket World

MPLS networks consist of LSRs connected via circuits called label switched paths (LSPs). To establish an LSP, a signaling protocol is required. Between two adjacent LSRs, an LSP is locally identified by a short, fixed-length identifier called a *label*, which is only significant

JUNIPER exhibit 1003

App. 6, pg. 178

between these two LSRs. When a packet enters an MPLS-based packet network, it is classified according to its forwarding equivalency class and, possibly, additional rules, which together determine the LSP along which the packet must be sent. For this purpose, the ingress LSR attaches an appropriate label to the packet and forwards the packet to the next hop. The label itself is a shim layer header, a virtual path identifier/virtual channel identifier (VPI/VCI) for ATM, or a data-link connection identifier (DLCI) for Frame Relay. When a packet reaches a core packet LSR, that LSR uses the label as an index into a forwarding table to determine the next hop, and the corresponding outgoing label. The LSR then writes the new label into the packet and forwards the packet to the next hop. When the packet reaches the egress LSR (or the one node before the egress LSR for penultimate hop popping), the label is removed and the packet is forwarded using appropriate forwarding, such as normal IP forwarding.

So how do these concepts apply to networks that are not packet-oriented, such as TDM- and WDM-based networks?

In TDM networks, the concept of label switching happens at the circuit level or segment level. Switching can happen, for example, at the time-slot level where an input OC3 time slot is cross-connected to an output OC3 time slot.

For WDM-capable nodes, switching happens at the wavelength level, where an input wavelength is cross-connected to an output wavelength. As such, SONET/SDH add/drop multiplexers (ADMs) and OXCs become equivalent to MPLS LSRs, time-slot LSPs and lambda LSPs become equivalent to packet-based LSPs, and the selection of time slots and wavelength becomes equivalent to the selection of packet labels. Also, nonpacket LSPs are bidirectional in nature, in contrast to packet LSPs, which are unidirectional (this is covered in more depth in Chapter 8).

The following section takes a closer look at label switching in TDM-based networks and touches upon label switching in WDM networks. The concepts of label switching in both TDM and WDM networks are similar in the sense that with TDM networks GMPLS controls circuits and with WDM GMPLS controls wavelengths.

## Label Switching in TDM Networks

SONET and SDH are two TDM standards that are used to multiplex multiple tributary signals over optical links, thus creating a multiplex structure called the *SONET/SDH multiplex*. Details about the SONET/SDH structure are covered in Appendix A, “SONET/SDH Basic Framing and Concatenation.”

If you choose to use the GMPLS control plane to control the SONET/SDH multiplex, you must decide which of the different components of the SONET/SDH multiplex that can be switched need to be controlled using GMPLS. As described in Appendix A, the SONET/SDH

frame format consists of overhead bytes, a payload, and a pointer to the payload. Essentially, every SONET/SDH element that is referenced by a pointer can be switched. These component signals in the SONET case are the synchronous transport signal (STS), Synchronous Payload Envelopes (SPEs), and virtual tributaries (VTs), such as STS-1, VT-6, VT-3, VT-2, and VT-1.5. For SDH, the elements that can be switched are the VC-4, VC-3, VC-2, VC-12, and VC-11.

When concatenation is used in the case of SONET or SDH, the new structure can also be referenced and switched using GMPLS. As explained in Chapter 2, “Metro Technologies,” concatenation—standard or virtual—allows multiple tributaries or STS/STM to be bonded to create a bigger pipe. GMPLS can be applied on the concatenated pipe.

The following sections discuss in more detail the concepts of label switching in a TDM network, including the following:

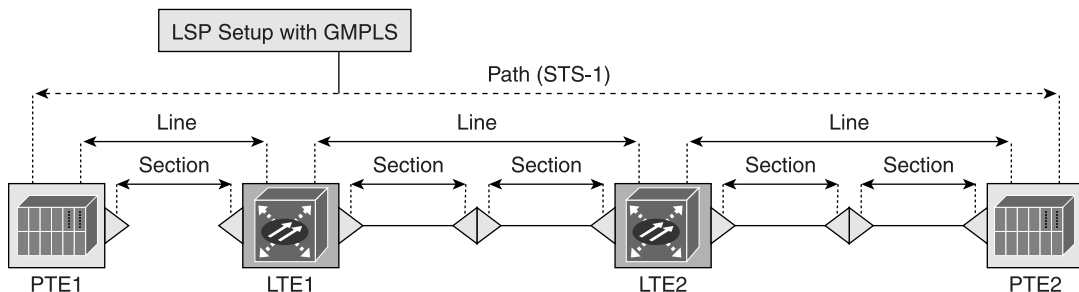
- Signaling in a TDM network
- SONET/SDH LSRs and LSPs
- The mechanics and function of a TDM label

## Signaling in a TDM Network

To support signaling in the TDM network, several modifications need to be made to MPLS. First, the traditional MPLS label needs to be modified to provide better binding between the label itself and the circuit it represents on a particular interface. Second, an LSP hierarchy needs to be introduced so that LSPs that represent signals can be tunneled inside other LSPs. Third, the capabilities of the label distribution protocols need to be extended so that they can distribute the information that is necessary to switch the signals along the path. A high-level description of the signaling modifications is covered in the next section, and a more detailed description is available in Chapter 8.

## SONET/SDH LSRs and LSPs

GMPLS defines a SONET/SDH terminal multiplexer, an ADM, and a SONET cross-connect as SONET/SDH LSRs. A path or circuit between two SONET/SDH LSRs becomes an LSP. A SONET/SDH LSP is a logical connection between the point at which a tributary signal (client layer) is adapted to its SPE for SONET or to its virtual container for SDH, and the point at which it is extracted from its SPE or virtual container. Figure 7-7 shows a SONET/SDH LSP. In this example, an STS-1 LSP is formed between path terminal equipment—PTE1 and PTE2—across line terminal equipment—LTE1 and LTE2. The LTEs are the SONET/SDH network elements that originate or terminate the line signal. The PTEs are the SONET/SDH network elements that multiplex/demultiplex the payload. A PTE, for example, would take multiple DS1s to form an STS-1 payload.

**Figure 7-7** GMPLS LSP Across SONET Equipment

To establish a SONET/SDH LSP, a signaling protocol is required to configure the input interface, switch fabric, and output interface of each SONET/SDH LSR along the path. A SONET/SDH LSP can be point-to-point or point-to-multipoint, but not multipoint-to-point, because no merging is possible with SONET/SDH signals. To facilitate the signaling and setup of SONET/SDH circuits, a SONET/SDH LSR must identify each possible signal individually per interface, because each signal corresponds to a potential LSP that can be established through the SONET/SDH LSR. GMPLS switching does not apply to all possible SONET/SDH signals—only to those signals that can be referenced by a SONET/SDH pointer, such as the STS SPEs and VTs for SONET and the VC-Xs for SDH.

The next section addresses the mechanics and functions of a GMPLS label in the context of TDM networks.

### The Mechanics and Function of a TDM Label

You have already seen label switching adopted with an asynchronous technology such as IP where a label attaches to an IP packet and helps put that packet on the right LSP in the direction of its destination. For SONET/SDH, which are synchronous technologies that define a multiplexing structure, GMPLS switching does not apply to individual SONET/SDH frames. GMPLS switching applies to signals, which are continuous sequences of time slots that appear in a SONET/SDH frame. GMPLS can switch SONET/SDH signals. As such, a SONET/SDH label needs to indicate the signals that can be switched, such as the STS SPE, VTs, and virtual containers.

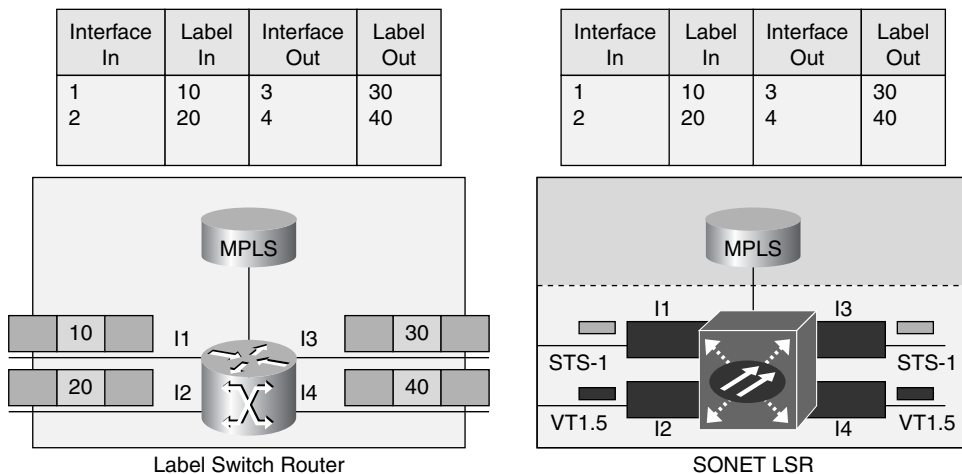
Figure 7-8 compares label switching applied to TDM and traditional label switching in the packet world.

As you can see, with a packet LSR, the labels are identified for a certain forwarding equivalency class and are used to label-switch the packet to its destination. The labels themselves are carried inside the IP packets for the LSR to perform the label-switching function. In the case of a



SONET/SDH LSR, the GMPLS control plane needs to map labels for the signals that need to be switched on each interface. In this example, the STS-1 signal on interface I1 is mapped to label 10 and is cross-connected to the STS-1 signal on interface I3, which is mapped to label 30. The VT 1.5 signal on interface I2 is mapped to label 20 and is cross-connected to the VT 1.5 signal on interface I4, which is mapped to label 40. Note that the SONET/SDH frames themselves do not carry any label; the mapping is just an indication by the GMPLS plane to allow the SONET/SDH node to perform the required switching function of the appropriate signals.

**Figure 7-8** SONET/SDH Label Switching



A SONET/SDH LSR has to identify each possible signal individually per interface to fulfill the GMPLS operations. To stay transparent, the LSR obviously should not touch the SONET/SDH overheads; this is why an explicit label is not encoded in the SONET/SDH overheads. Rather, a label is associated with each individual signal and is locally unique for each signal at each interface.

Because the GMPLS label is not coded in the signal itself, a mechanism needs to be established to allow the association of a label with SONET/SDH signals. The GMPLS label is defined in a way that enables it to give information about the SONET/SDH multiplex, such as information about the particular signal and its type and position in the multiplex.

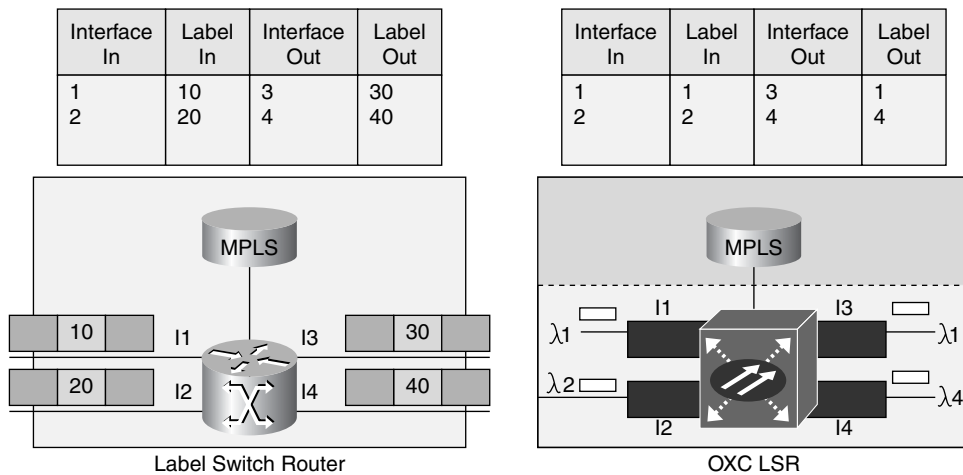
## Label Switching in WDM Networks

WDM is a technology that allows multiple optical signals operating at different wavelengths to be multiplexed onto a single fiber so that they can be transported in parallel through the fiber. OXCs in turn cross-connect the different wavelengths, in essence creating an optical path from

source to destination. The optical path itself can carry different types of traffic, such as SONET/SDH, Ethernet, ATM, and so on. OXCs can be all optical, cross-connecting the wavelengths in the optical domain, or they can have optical-to-electrical-to-optical conversion, which allows for wavelength conversion mechanisms. In the GMPLS context, OXCs would run the GMPLS control plane and would become comparable to LSRs. Lambda LSPs are considered similar to packet-based LSPs, and the selection of wavelengths and OXC ports is considered similar to label selection.

Figure 7-9 compares the concept of MPLS switching in a WDM network in the same way that Figure 7-8 did for the TDM network.

**Figure 7-9** WDM Label Switching





As already shown in the TDM example, the GMPLS labels are not carried inside the actual packet. In the case of an OXC LSR, the GMPLS control plane needs to map labels for the lambdas that need to be switched on each interface. In this example, label 1 on interface I1 is mapped to lambda 1 and cross-connected to lambda 1 on I3, which is mapped to label 1. Again, because the GMPLS label is not coded in the wavelength, a mechanism needs to be established to associate lambdas with labels. This is discussed in Chapter 8.

## Conclusion

As discussed in this chapter, GMPLS is necessary to establish a dynamic way to provision optical networks. You have seen the benefits and drawbacks of both the static centralized and dynamic decentralized provisioning models. The chapter also discussed the different signaling models,

such as the overlay, peer, and augmented models. These resemble how IP packet-based networks are deployed today over ATM or Frame Relay circuit-based networks. You have also seen how GMPLS uses labels to cross-connect the circuits for TDM and WDM networks. Although the concept of labels was adopted, the use of these labels is quite different from the traditional use of labels in data forwarding.

The next chapter goes into more detail about the extensions to routing and signaling that were added to the traditional MPLS control plane to accommodate optical networks.



This chapter covers the following topics:

- GMPLS Interfaces
- Modification of Routing and Signaling
- Inclusion of Technology-Specific Parameters
- Link Management Protocol
- GMPLS Protection and Restoration Mechanisms
- Summary of Differences Between MPLS and GMPLS

# GMPLS Architecture

---

Optical networks present some added challenges that do not normally exist in packet-switched networks (PSNs) and hence cannot be fully addressed by the traditional MPLS schemes. Here are a few examples of these challenges:

- Optical/TDM bandwidth allocation is done in discrete amounts, whereas in PSNs, bandwidth can be allocated from a continuous spectrum.
- The number of links in an optical network can be orders of magnitude larger than in a traditional network, due to the possible explosion in the number of parallel fibers deployed and the number of lambdas on each fiber. This in turn raises the issues of IP address assignment for optical links and the manageability of connecting ports on different network elements. If a fiber has 32 wavelengths, for example, between points A and B, and if each wavelength is treated as a separate link with its own addressing, the one fiber will create 32 different networks that need to be addressed and managed.
- Fast fault detection and isolation have always been advantages that optical networks have over PSNs.
- The fact that user data in an optical network is transparently switched necessitates the decoupling of user data from control plane information.

Generalized MPLS (GMPLS) attempts to address these challenges by building on MPLS and extending its control parameters to handle the scalability and manageability aspects of optical networks. This chapter explains the characteristics of the GMPLS architecture, such as the extensions to routing and signaling and the addition of technology parameters, that GMPLS adds to MPLS to be able to control optical networks.

## GMPLS Interfaces

The GMPLS architecture extends MPLS to include five different types of interfaces used on label switch routers:

- **Packet-switch capable (PSC) interfaces**—Interfaces that can recognize packet boundaries and forward data based on packet headers. This is typical of interfaces on routers and L3 Ethernet switches.

- **Layer 2–switch capable (L2SC) interfaces**—Interfaces that can recognize L2 cell or frame boundaries and forward data based on L2 headers. This is typical of interfaces on ATM switches, Frame Relay switches, and L2 Ethernet switches.
- **Time-division multiplexing (TDM) interfaces**—Interfaces that can recognize time slots and forward data based on the data’s time slot in a repeating cycle. This is typical of interfaces on digital cross-connects (DACs), SONET add/drop multiplexers (ADMs), and SONET cross-connects. Such interfaces are referred to as TDM capable.
- **Lambda-switch capable (LSC) interfaces**—Interfaces that can forward data based on the lambda (wavelength) it was received on. This is typical of optical cross-connects (OXCs) that switch traffic on the wavelength level.
- **Fiber-switch capable (FSC) interfaces**—Interfaces that can forward data based on the position of the data in real-world physical spaces. This is typical of OXCs that switch traffic on the fiber or multiple-fiber level.

## Modification of Routing and Signaling

The development of GMPLS requires the modification of current routing and signaling protocols. The adoption of a common, standardized control plane for managing packet/cell switches and optical switches is extremely important to the networking industry. This introduces a unified method for achieving fast provisioning, restoration, routing, monitoring, and managing data-switched and optical-switched networks while maintaining interoperability between multiple vendors. The MPLS control plane is being extended from controlling data switches to a more generic role of controlling any type of switching, including optical switching—hence the term Generalized MPLS.

To help MPLS span switches that are not packet-oriented, GMPLS introduces some modifications to MPLS in the areas of routing and signaling. The modifications take place in the following areas:

- Enhancements to routing protocols
- Enhancements to signaling protocols

The following sections discuss routing and signaling enhancements.

## Enhancements to Routing

Introducing routing into TDM and optical networks does not mean turning TDM and optical nodes into IP routers, but rather using the benefits of routing protocols as far as relaying paths and resource information to better use network resources. In optical and TDM networks, this information includes the following:

- The available capacity of the network links
- The switching and termination capabilities of the nodes and interfaces
- The link’s protection properties

This information is carried inside routing protocols such as Open Shortest Path First for Traffic Engineering (OSPF-TE) and IS-IS Traffic Engineering (IS-IS-TE). GMPLS introduces extensions to OSPF-TE and IS-IS-TE to allow these protocols to tailor to the specific information required by these networks. OSPF-TE and IS-IS-TE are extensions of the OSPF and IS-IS routing protocols that allow them to carry network information about available network resources. This information is used by protocols such as RSVP-TE to engineer the traffic in the network.

An MPLS TE link is considered to be like any regular link, meaning a link where a routing protocol adjacency is brought up via protocols such as OSPF. The link's Shortest Path First (SPF) properties and the TE properties are calculated and advertised. For GMPLS to accommodate optical networks, a few variations need to be introduced:

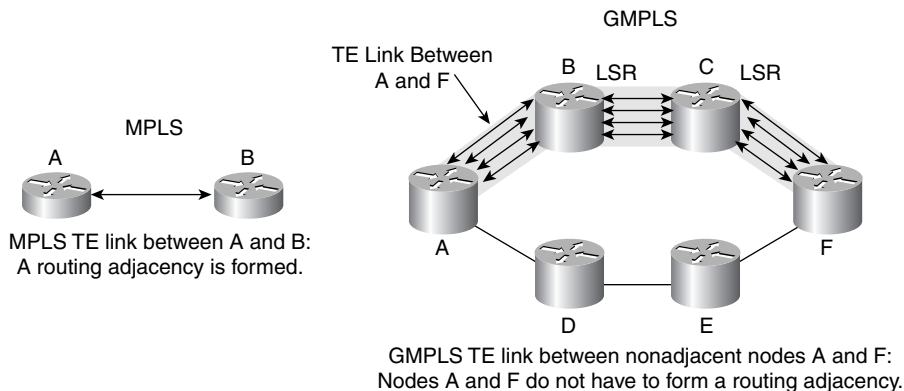
- Nonpacket links can be brought up without establishing a routing adjacency.
- A label switched path (LSP) can be advertised as a point-to-point TE link, and the advertised TE link need no longer be between two OSPF/IS-IS direct neighbors.
- A number of links can be advertised as a single TE link, and there is no one-to-one association between routing adjacencies and a TE link.

A GMPLS TE link has special TE properties that can be configured or obtained via a routing protocol. An example of TE properties would be the bandwidth accounting for the TE link, including the unreserved bandwidth, the maximum reservable bandwidth, and the maximum LSP bandwidth. Other properties include protection and restoration characteristics.

IS-IS-TE and OSPF-TE explain how to associate TE properties to regular (packet-switched) links. GMPLS extends the set of TE properties and also explains how to associate TE properties with links that are not packet-switched, such as links between OXC's.

Figure 8-1 shows a TE link.

**Figure 8-1** GMPLS TE Link



As shown in Figure 8-1, a GMPLS TE link extends beyond two adjacent nodes and can include multiple parallel component links. The end nodes of the link do not have to be part of a routing adjacency. In the context of MPLS, the link is between two adjacent nodes A and B and forms a routing adjacency using a routing protocol, say OSPF. In the GMPLS context, the link traverses multiple nodes and the two label switch routers (LSRs) B and C. A and F do not have to establish a routing adjacency.

The GMPLS enhancements to routing include the following:

- LSP hierarchy—routing
- Unnumbered links
- Link bundling
- Link protection types
- Shared link group information
- Interface switching capability descriptor

The next sections examine each of these enhancements to routing introduced by GMPLS.

## LSP Hierarchy—Routing

The difference between the traditional fiber networks and WDM networks is that WDM introduces a significant increase in the number of paths between two endpoints, mainly because it introduces hundreds of wavelengths on each fiber. Couple that with the possibility of tens and hundreds of fibers between two optical switches, and the number of paths could become challenging to traditional routing protocols if every path (LSP) is considered a separate link in interior routing protocols such as OSPF and IS-IS.

LSP hierarchy can address this issue by allowing LSPs to be aggregated inside other LSPs. There is a natural order for this aggregation that is based on the multiplexing capability of the LSP types. With GMPLS, LSPs start and end on devices of the same kind, such as routers, TDM switches, WDM switches, and fiber switches. An LSP that starts and ends on a packet-switch-capable (PSC) interface can be nested with other LSPs into an LSP of type TDM that starts and ends on a TDM interface, which can be nested in LSC-LSPs that start and end on an LSC interface, which could be nested in FSC-LSPs that start and end on FSC interfaces. This is illustrated in Figure 8-2.

When an LSR establishes an LSP, it can advertise the LSP in its instance of routing protocol (OSPF or IS-IS) as a TE link. This link is called a *forwarding adjacency (FA)*. The LSP itself is referred to as the forwarding adjacency LSP, or FA-LSP.

IS-IS/OSPF floods the information about FAs just as it floods the information about any other links. As a result of this flooding, an LSR has in its TE link-state database information about not just basic TE links, but FAs as well. Figure 8-2 shows how GMPLS FA-LSP can be carried within other FA-LSPs. The different FA-LSPs introduced in this figure are FA-LSCs, FA-TDMs, and FA-PSCs.



Figure 8-2 GMPLS LSP Hierarchy

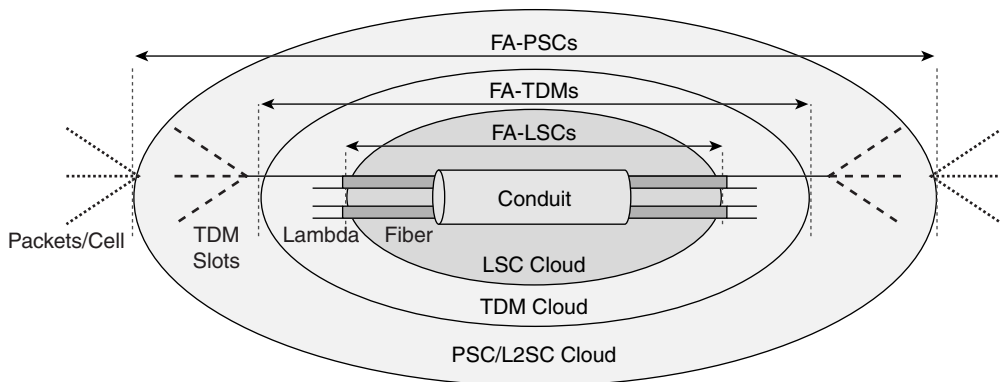


Figure 8-2 shows the following:

- 1 FA-LSCs are formed by nodes that sit at the boundary of a lambda cloud and a fiber cloud. The FA-LSCs get advertised in the routing protocols and are available to be used as any other TE links.
- 2 Nodes that sit at the boundary of a TDM cloud and a lambda cloud form FA-TDMs. The FA-TDMs get advertised as TE links.
- 3 Nodes that sit at the boundary of a PSC/L2SC and TDM cloud form FA-PSCs or FA-L2SCs that get advertised as TE links.
- 4 Low-order packet LSPs can be combined and tunneled inside higher-order FA-PSCs. In the same manner, low-order FA-PSCs can be combined and tunneled inside higher-order FA-TDMs, which can be combined and tunneled inside higher-order FA-LSCs.
- 5 FAs (links) are either numbered or unnumbered and can be bundled according to the GMPLS bundling procedures.

## Unnumbered Links

As in an IP network, the nodes in an optical network have to be addressed and referenced. Addressing these nodes helps identify not only the nodes but also the components—that is, the links of each of these nodes. Addressing allows signaling protocols such as RSVP to establish optical paths across the OXCs.

In normal routing, each link in the network can be identified via its own subnet. This has proven to be challenging even in packet networks because it requires the assignment and management of many small subnets. In optical networks, in which the number of links can increase dramatically, IP address assignment proves much more challenging because a fiber can carry hundreds of wavelengths. Thus, the concept of unnumbered links should be quite useful.

An *unnumbered link* is a point-to-point link that is referenced using a link identifier. The link identifier is a unique, nonzero, 32-bit local identifier. The identifier for the local node is called the *local link identifier*, while the link identifier for the remote node is called the *remote link identifier*. If the remote link identifier is not known, a 0 identifier is used instead.

A network node can be addressed via a router ID (normally the highest or lowest IP address on that node). The links on that node can then be identified locally via the tuple (router ID, link number). Exchanging the identifiers may be accomplished by multiple methods, including configuration, LMP, RSVP-TE, IS-IS/OSPF, and so on.

Figure 8-3 illustrates the concept of unnumbered links.

**Figure 8-3** *Unnumbered Links*

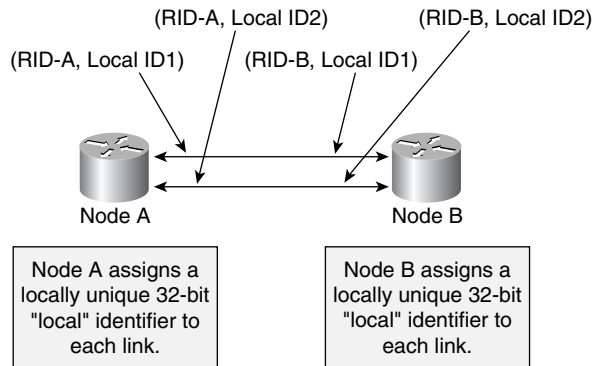


Figure 8-3 shows how node A identifies each link with a tuple formed with its router ID RID-A and the local link identifier.

Current signaling used by MPLS TE doesn't provide support for unnumbered links because the current signaling doesn't provide a way to indicate an unnumbered link in its EXPLICIT\_ROUTE object (ERO) and RECORD\_ROUTE object (RRO). Extensions to RSVP-TE define an optional object called LSP\_TUNNEL\_INTERFACE\_ID that could be used in RSVP PATH or Reservation (RESV) messages. The LSP\_TUNNEL\_INTERFACE\_ID object is an LSR router ID and a 32-bit interface ID tuple. Also, subobjects of the ERO and RRO are defined for the support of unnumbered links.

### Link Bundling

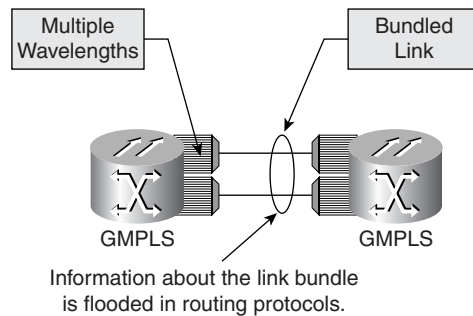
Link bundling allows multiple TE links to be bundled into one bigger TE link. The subset links are called *component links*, and the group of links is called a *bundled link*.

On a bundled link, a combination of <(bundled) link identifier, component link identifier, label> is sufficient to unambiguously identify the appropriate resources used by an LSP.

Link bundling improves routing by reducing the number of links and associated attributes that are flooded into routing protocols such as OSPF and IS-IS. Link bundling allows multiple parallel

links of similar characteristics to be aggregated and flooded as a bundled link. Figure 8-4 shows this concept.

**Figure 8-4** *Link Bundling*



All component links in a bundle must:

- Begin and end on the same pair of LSRs
- Have the same link type, such as point-to-point or multiaccess
- Have the same TE metric
- Have the same set of resource classes at each end of the links

A bundled link is considered alive if one of its component links is alive. Determining the liveness of the component links can be done via routing protocols, LMP, or L1 or L2 information. Once a bundled link is considered alive, the information about the bundled link is flooded as a TE link.

---

**WARNING** The benefits of link bundling in reducing the number of flooded links come at the expense of loss of information. Link bundling involves the aggregation of the component links, and in the process of summarizing the attributes of several links into a bundled link, information is lost. Remember that the information that is flooded in the routing protocols is information about the bundled link itself and *not* information about the component links. As an example, when multiple parallel SONET links are summarized, information about the total reservable bandwidth of the component links is advertised, but information about the bandwidth and time slots of each link is lost.

---

While the link-state protocols carry a single bundled link, signaling requires that individual component links be identified. Because the ERO does not carry information about the component links, the component link selection becomes a local matter between the LSR bundle neighbors. LMP offers a way to identify individual component links. (LMP is described later in the chapter, in the section “Link Management Protocol.”)

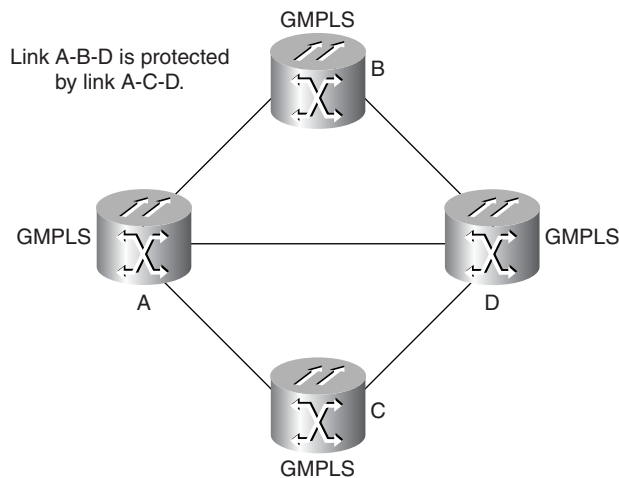
## Link Protection Types

GMPLS introduces the concept of a *link protection type*, which indicates the protection capabilities that exist for a link. Path computation algorithms use this information to establish links with the appropriate protection characteristics. This information is organized in a hierarchy where typically the minimum acceptable protection is specified at path instantiation and a path selection technique is used to find a path that satisfies at least the minimum acceptable protection. The different link protection types are as follows:

- **Extra Traffic**—This type of link protects another link or links. In case of failure of the protected links, all LSPs on this link are lost.
- **Unprotected**—This type of link is simply not protected by any other link. If the unprotected link fails, all LSPs on the link are lost.
- **Shared**—This type of link is protected by one or more disjoint links of type Extra Traffic.
- **Dedicated 1:1**—This type of link is protected by a disjoint link of type Extra Traffic.
- **Dedicated 1+1**—This type of link is protected by a disjoint link of type Extra Traffic. However, the protecting link is not advertised in the link-state database and therefore is not used by any routing LSPs.
- **Enhanced**—This type of link indicates that a protection scheme that is more reliable than Dedicated 1+1 should be used—for example, four-fiber BLSR.

Figure 8-5 shows the different protection types.

**Figure 8-5** *Link Protection Types*



Link A-B-D is protected by link A-C-D. Link A-C-D is of type Extra Shared. The following protection scenarios can occur:

- **Link A-B-D is 1+1 protected**—Link A-C-D protects link A-B-D. Link A-C-D is not advertised and hence does not carry any LSPs unless link A-B-D fails.

- **Link A-B-D is 1:1 protected**—Link A-C-D protects link A-B-D. Link A-C-D is advertised and can carry LSPs, but it gets preempted to protect link A-B-D if link A-B-D fails.

### Shared Risk Link Group Information

A set of links may constitute a *shared risk link group (SRLG)* if they share a resource whose failure may affect all links in the set. Multiple fibers in the same conduit, for example, could constitute an SRLG because a conduit cut may affect all the fibers. The same applies to multiple lambdas in a fiber that can all be affected if a fiber cut occurs. The SRLG is an optional 32-bit number that is unique within an IGP domain. A link might belong to multiple SRLGs. The SRLG of an LSP is the union of the SRLGs of the links in the LSP. The SRLG information is used to make sure that diversely routed LSPs do not have a common SRLG—that is, they do not share the same risks of failure. Figure 8-6 illustrates the concept of an SRLG.

**Figure 8-6** Shared Risk Link Group

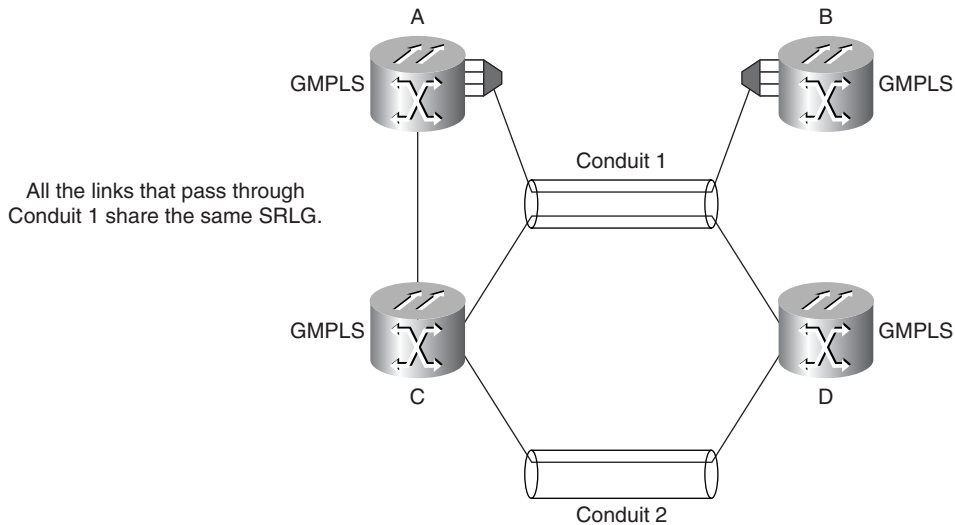


Figure 8-6 shows that all links that pass through conduit 1 share the same SRLG. The same is true for all links that pass through conduit 2. If the SRLG option is used, two LSPs that need to be diversely routed between node A and node D cannot both pass through conduit 1 or conduit 2, because they would have the same SRLGs in common.

### Interface Switching Capability Descriptor

In the context of GMPLS, a link is connected to a node via an interface. An interface on the same node and on either side of the link may have multiple switching capabilities. The interface switching capability descriptor is used to handle interfaces that support multiple switching capabilities, for interfaces that have Max LSP Bandwidth values that differ by priority level (P), and for interfaces that support discrete bandwidth. A fiber interface, for example, that is connected to a node can carry multiple lambdas, and each lambda can be

terminated. If the lambda is carrying packets, packet-switching can be performed. If the lambda is carrying a TDM circuit, the TDM circuit is switched. If the lambda is not terminated at the node, the lambda itself can be lambda switched. To support such interfaces, a link-state advertisement would carry a list of interface switching descriptors.

You saw in the “GMPLS Interfaces” section that GMPLS defines five types of interfaces: PSC, L2SC, TDM, LSC, and FSC. The following list describes the interface descriptors associated with these types of interfaces:

- For the PSC interfaces, various levels of PSC from 1 through 4 exist to establish a hierarchy of LSPs tunneled within LSPs, with PSC 1 being the highest order.
- For interfaces of type PSC1 through 4, TDM, and LSC, the interface descriptor carries additional information in the following manner:
  - For PSC interfaces, the additional information includes Maximum (Max) LSP Bandwidth, Minimum (Min) LSP Bandwidth, and interface MTU.
  - For TDM-capable interfaces, the additional information includes Maximum LSP Bandwidth, information on whether the interface supports standard or arbitrary SONET/SDH, and Minimum LSP Bandwidth.
  - For LSC interfaces, the additional information includes Reservable Bandwidth per priority, which specifies the bandwidth of an LSP that can be supported by the interface at a given priority number.

### Determining the Link Capability

The link capability is determined based on the tuple <interface switching capability, label>. Carrying label information on a given TE link depends on the interface switching capability at both ends of the link and is determined as follows:

- [PSC, PSC]—The label is carried in the “shim” header (RFC 3032, *MPLS Label Stack Encoding*).
- [TDM, TDM]—The label represents a TDM time slot.
- [LSC, LSC]—The label represents a port on an OXC.
- [PSC, TDM]—The label represents a TDM time slot.
- [TDM, LSC]—The label represents a port.

### Interface Switching Capability Descriptor Examples

The following are examples of interface switching capability descriptors.

Fast Ethernet 100-Mbps Ethernet packet interface on an LSR:

- Interface switching capability descriptor:
  - Interface Switching Capability = PSC-1
  - Encoding = Ethernet 802.3

- Max LSP Bandwidth[P] = 100 Mbps for all P (where P indicates the LSP priority level; a priority of 7, for example, gives the LSP high priority)

The following is how the interface descriptor is represented for an OC-192 SONET interface on a digital cross-connect with Standard SONET.

Assuming that it is possible to establish the following connections, VT-1.5, STS-1, STS-3c, STS-12c, STS-48c, STS-192c, the interface switching capability descriptor of that interface can be advertised as follows:

- Interface Switching Capability = TDM [Standard SONET]
- Encoding = SONET ANSI T1.105
- Min LSP Bandwidth = VT1.5
- Max LSP Bandwidth[p] = STS192 for all p (where p refers to LSP priority)

## Enhancements to Signaling

GMPLS enhances the traditional MPLS control plane to support additional different classes of interfaces, such as TDM, LSC, and FSC. The support of these interfaces requires some changes to signaling, such as the following:

- LSP hierarchy—signaling
- Enhancements to labels
- Bandwidth encoding
- Bidirectional LSPs
- Notification of label error
- Explicit label control
- Protection information
- Administrative status information
- Separation of control and data channels
- Notify messages

The following sections describe the different enhancements to signaling introduced by GMPLS.

### LSP Hierarchy—Signaling

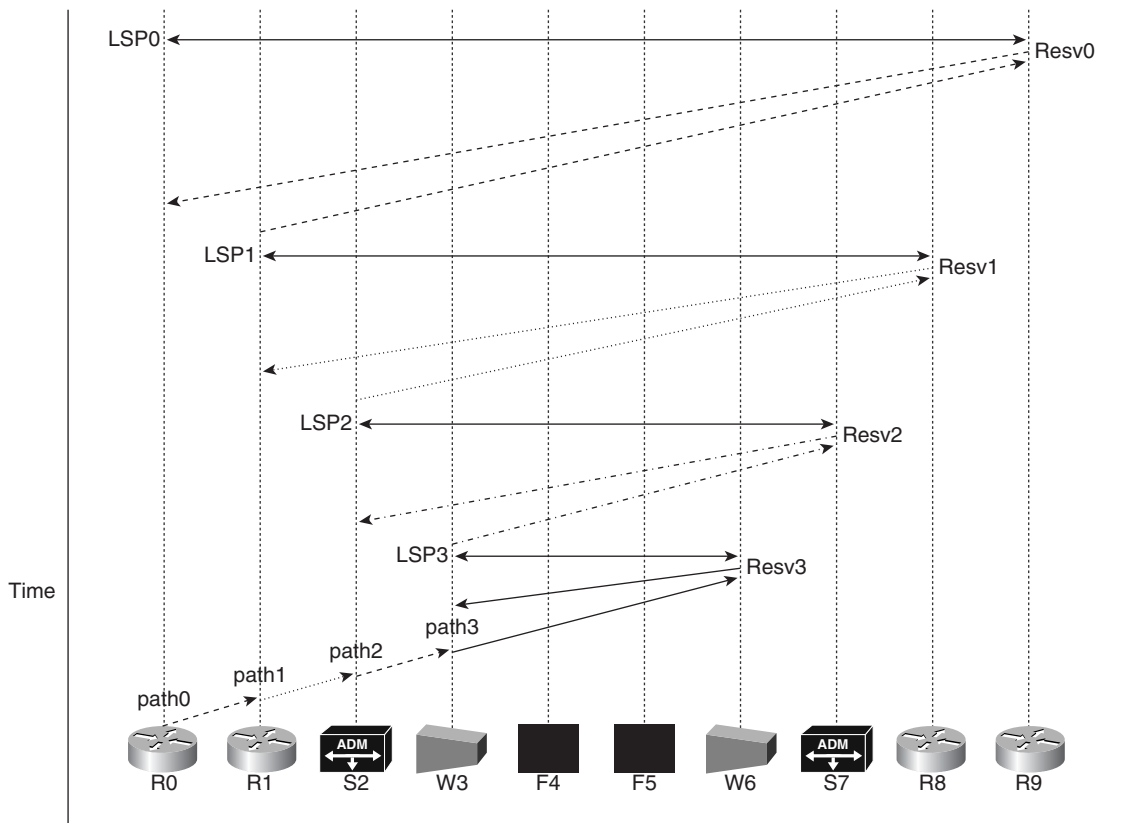
As already explained in the “LSP Hierarchy—Routing” section, GMPLS supports the concept of hierarchical LSPs, which allows multiple LSPs to be nested; that is, it allows newly initiated LSPs to be aggregated within existing LSPs. The newly initiated LSPs are tunneled inside an existing higher-order LSP, which becomes a link along the path of the new LSP. This dramatically enhances network scalability and manageability because it minimizes the number of elements that are flooded and advertised within the network. This section explains the signaling aspect of the LSP hierarchy.

To give an example of how GMPLS signaling uses the LSP hierarchy, assume that a certain router requests bandwidth to be allocated along a network consisting of data switches, SONET cross-connects, WDM-capable switches, and fiber switches.

The request from the edge router to establish a PSC LSP with a certain bandwidth could trigger the establishment of multiple higher-order LSPs that get initiated by other switches along the path. Lower-order LSPs (the new LSPs) get nested inside the higher-order LSPs that already exist or that get triggered based on the edge router's request.

Figure 8-7 shows the establishment of a series of LSPs along a path that consists of routers (R0, R1, R8, and R9), SONET ADMs (S2 and S7), WDM Optical Electrical Optical (OEO) switches (W3 and W6), and fiber switches (F4 and F5). A PATH request, path 0, needed for the formation of LSP0 between R0 and R9, is sent from R0 to R1. At router R1, this triggers the initiation of LSP1 between R1 and R8. LSP1 is nested inside LSP0. The PATH messages—path1, path2, and path3—continue to propagate, and the LSPs keep getting created until the final establishment of LSP0 between R0 and R9.

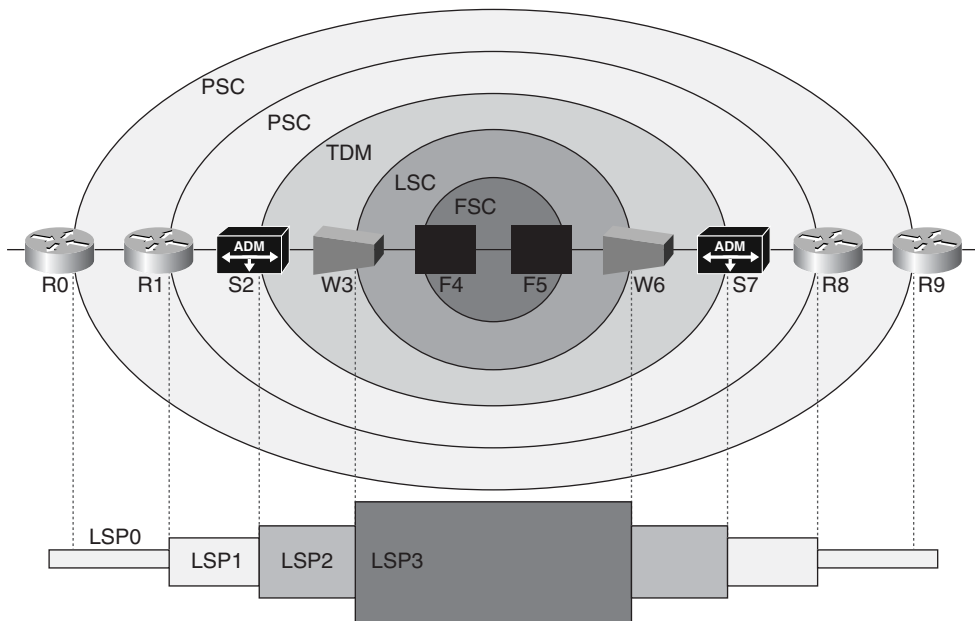
Figure 8-7 Initiation of New Nested LSPs





An LSP is established when the path message has completed its path inside higher-level LSPs and a RESV message is received. Note in Figure 8-8 how LSP3, the higher-level LSP, gets established first, then LSP2 gets established inside LSP3, then LSP 1 inside LSP2, and LSP 0 inside LSP1.

**Figure 8-8** *Nested LSPs*



Now assume that a carrier is offering an Ethernet packet transport service between two service providers—ISP1 and ISP2—with an SLA set to 200 Mbps. For simplicity, assume that the carrier's end-to-end network is formed via routers (R0, R1, R8, and R9), SONET ADMs (S2 and S7), WDM OEO switches (W3 and W6), and fiber switches (F4 and F5). Also, for the sake of simplicity, the GE service for the carrier is assumed to be point-to-point between R0 and R9, meaning that all traffic that comes in on the GE links of R0 comes out on the GE links of R9. Physical connectivity is done in the following way:

- **R0-R1 and R8-R9**—Ethernet GE (1 Gbps) link
- **R1-S2 and R8-S7**—OC48c (2.4 Gbps) packet over SONET (PoS) link
- **S2-W3 and S7-W6**—OC192 (9.6 Gbps) TDM link
- **W3-F4 and W6-F5**—16 OC192 lambdas
- **F4-F5**—16 fibers, carrying 16 OC192 lambdas each

The following illustrates the process of LSP creation on all the boxes between ISP1 and ISP2:

- LSP0 between R0 and R9 as a 200-Mbps connection
- LSP1 between R1 and R8 as an OC48c connection
- LSP2 between S2 and S7 as an OC192 connection
- LSP3 between W3 and W6 as a lambda connection
- LSP4 between P4 and P5 as a fiber connection

LSP0 is nested inside LSP1, LSP1 is nested inside LSP2, and LSP2 is nested inside LSP3.

In addition to the creation of the LSPs, the nodes announce the residual bandwidth available in the LSP hierarchy in the following manner:

- 1 Node R0 announces a PSC link from R0 to R9 with bandwidth equal to the difference between the GE link and 200 Mbps—that is, 800 Mbps.
- 2 Node R1 announces a PSC link from R1 to R8 with bandwidth equal to the difference between the OC48c capacity (2.4 Gbps) and 200 Mbps—that is, 2.2 Gbps.
- 3 Node S2 announces a TDM link from S2 to S7 with bandwidth equal to the difference between the OC192 (STS-192) link capacity and the allocated OC48 (STS-48) time slots—that is, STS-144.
- 4 Node W3 announces an LSC link from W3 to W6 with bandwidth equal to the difference between 16 lambdas and the allocated lambda—that is, 15 lambdas.
- 5 Node P4 announces an FSC link from P4 to P5 with bandwidth equal to the difference between 16 fibers and the allocated fiber—that is, 15 fibers.

As part of enhancements to signaling, GMPLS introduces enhancements to the MPLS label itself, as described next.

## Enhancements to Labels

GMPLS introduces new label concepts to accommodate the specific requirements of the optical space. The new concepts include the generalized label, the label set, and the suggested label.

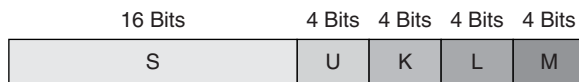
### The Generalized Label

To accommodate the scope of GMPLS that includes non-packet/cell interfaces, several new forms of labels are required, which are called *generalized labels*. A generalized label extends the traditional label by allowing the label to identify time slots, wavelengths, or space-division multiplexed positions. Examples are label representation of a fiber in a bundle, a waveband within a fiber, a wavelength in a waveband, and a set of time slots within a wavelength, as well as the traditional MPLS label. The generalized label has enough information to allow the receiving node to program a cross-connect regardless of the type of the cross-connect. As you

have already seen in Chapter 7, “MPLS Controlling Optical Switches,” the label is purely a signaling construct used to give information about how interfaces are cross-connected and is not part of the forwarding plane.

An example of a SONET/SDH label format is shown in Figure 8-9. This is an extension of the (K, L, M) numbering scheme defined in ITU-T Recommendation G.707, “Network Node Interface for the Synchronous Digital Hierarchy” (October 2000). The S, U, K, L, and M fields help identify the signals in the SONET/SDH multiplex. Each letter indicates a possible branch number starting at the parent node in the SONET/SDH multiplex structure.

**Figure 8-9** SONET/SDH Label Format



A generalized label request is used to communicate the characteristics required to support the LSP being requested. The information carried in the generalized label request includes the following:

- **LSP Encoding Type**—An 8-bit field that indicates the LSP encoding types, such as packet, Ethernet, PDH, SDH, SONET, Digital Wrapper (DW), lambda, fiber, and Fiber Channel.

When a generalized label request is made, the request carries an LSP encoding type parameter that indicates the type of the LSP, such as SONET, SDH, Gigabit Ethernet, lambda, fiber, and so on. The lambda encoding type, for example, refers to an LSP that encompasses a whole wavelength. The fiber encoding type refers to an LSP that encompasses a whole fiber port. The encoding type represents the type of the LSP and not the nature of the links the LSP traverses. A link may support a set of encoding formats where the link can carry and switch a signal of one or more of these encoding formats depending on the link’s resource availability and capacity.
- **Switching Type**—An 8-bit field that indicates the type of switching that should be performed on a particular link. This field is needed for links that advertise more than one type of switching capability, such as PSC, L2SC, TDM, LSC, and FSC.
- **Generalized Payload Identifier (G-PID)**—A 16-bit field used by the nodes at the endpoint of the LSP to identify the payload carried by the LSP. Examples of the PID are standard Ethertype values for packet and Ethernet LSPs. Other values include payload types such as SONET, SDH, Digital Wrapper (DW), STS, POS, ATM mapping, and so on.

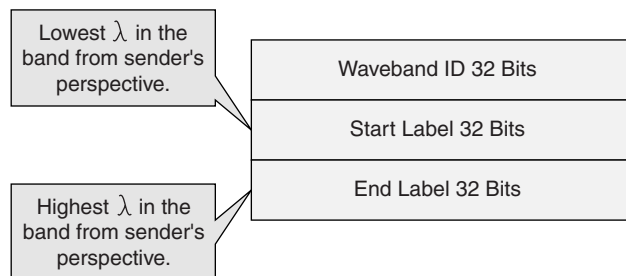
A generalized label carries only a single level of labels—that is, the label is nonhierarchical. When multiple levels of labels are required, each LSP must be established separately, as discussed in the previous section “LSP Hierarchy—Signaling.”

## Waveband Switching Support

A *waveband* represents a set of contiguous wavelengths that can be switched together to a new waveband. For optimization reasons, it may be desirable for an OXC to optically switch multiple wavelengths as a unit. This may reduce the distortion on the individual wavelengths and allow tighter separation of the individual wavelengths. The waveband label is defined to support this special case.

Waveband switching uses the same format as the generalized label. Figure 8-10 shows the format of the generalized label in the context of waveband switching.

**Figure 8-10** *Generalized Label—Waveband Switching*



## The Label Set

The *label set* is used to restrict the label ranges that may be used for a particular LSP between two peers. The receiver of a label set must restrict its choice to one label range that is in the label set. The label set is useful in the optical domain because of the restrictions on how optical equipment allocates wavelengths and handles wavelength conversion, which restricts the use of labels that are bound to these wavelengths. Reasons for using the label set include the following:

- The end equipment can transmit and receive only on a small, specific set of wavelengths/bands.
- There is a sequence of interfaces that cannot support wavelength conversion and that requires the same wavelength to be used end-to-end over a sequence of hops or an entire path.
- For operators, it is desirable to limit the amount of wavelength conversion being performed to reduce the distortion of the optical signals.
- The two ends of a link support different sets of wavelengths.

The use of a label set is optional, and if it is not present, it is assumed that all labels can be used.

## The Suggested Label

GMPLS allows an upstream node to suggest a label to the downstream (one hop away) node for different optimization purposes that are specific to optical networks. The downstream node may override the suggested label at the expense of higher LSP setup times and perhaps suboptimal allocation of network resources. A typical example is when an optical switch configures its own label to adjust its mirrors and save valuable time before the downstream switch allocates the label. Other examples involve any activity where there is latency in configuring the switching fabric.

Early configuration can reduce setup latency and may be important for restoration purposes where alternate LSPs may need to be rapidly established as a result of network failures.

## Bandwidth Encoding

GMPLS LSPs support packet or nonpacket LSPs. For nonpacket LSPs, it is useful to list the discrete bandwidth value of the LSP. Bandwidth encoding values include values for DS0 to OC768, E1 to STM-256, 10/100/1000/10,000-Mbps Ethernet, and 133- to 1062-Mbps Fiber Channel. The bandwidth encodings are carried in protocol-specific (RSVP-TE, CR-LDP) objects. Examples of RSVP-TE are the SENDER\_TEMPLATE and FLOW\_SPEC objects.

## Bidirectional LSPs

Many optical service providers consider bidirectional optical LSPs a requirement, because many of the underlying constructs for SONET/SDH networks are inherently bidirectional. It is assumed that bidirectional LSPs have the same TE requirements (including fate sharing, protection, and restoration) and resource requirements (such as latency and jitter) in each direction.

The traditional MPLS LSP establishment is unidirectional. Establishing a bidirectional LSP requires establishing two unidirectional LSPs, which has many disadvantages:

- The latency to establish the bidirectional LSP is equal to one round-trip signaling time plus one initiator-terminator signaling transit delay. This extends the setup latency for successful LSP establishment and extends the worst-case latency for discovering an unsuccessful LSP. These delays are particularly significant for LSPs that are established for restoration purposes.
- The control overhead of two unidirectional LSPs is twice that of one bidirectional LSP, because separate control messages must be generated for each unidirectional LSP.
- Because the resources are established in separate segments, route selection gets complicated. Also, if the resources needed to establish the LSP are not available, one unidirectional LSP gets established, but the other doesn't. This decreases the overall probability of successful establishment of the bidirectional connection.

- SONET equipment in particular relies on hop-by-hop paths for protection switching. SONET/SDH transmits control information in-band. This requires connections to be paired, meaning that bidirectional LSP setup is highly desirable. Therefore, GMPLS supports additional methods that allow bidirectional LSP setup, to reduce session establishment overhead.

### Notification of Label Error

Some situations in traditional MPLS and GMPLS result in an error message containing an “Unacceptable label value” indication. When these situations occur, it is useful if the node that is generating the error message indicates which labels are acceptable. To cover these situations, GMPLS introduces the ability to convey such information via an acceptable label set. An acceptable label set is carried in appropriate protocol-specific error messages.

The format of an acceptable label set is identical to a label set, as described earlier in this chapter in the section “The Label Set.”

### Explicit Label Control

As discussed in Chapter 7, with RSVP-TE, the interfaces used by an LSP may be controlled by an explicit route via the ERO or ERO hop. This allows the LSP to control which nodes/interfaces it goes in and out on. The problem is that the ERO and ERO hop do not support explicit label subobjects, which means that they cannot support the granularity needed by optical networks. For example, in networks that are not packet-based, LSPs sometimes need to be spliced together. This means that the tail end of an LSP needs to be spliced with the head end of another LSP. GMPLS introduces the ERO subobject/ERO hop to allow finer granularity for explicit routes.

### Protection Information

GMPLS uses a new object type length value (TLV) field to carry LSP protection information. The use of this information is optional. Protection information indicates the LSP’s link protection type. When a protection type is indicated, the connection request is processed only if the desired protection type can be honored. A link’s protection capabilities may be advertised in routing.

Protection information also indicates whether the LSP is a primary or secondary LSP. A secondary LSP is a backup to a primary LSP. The resources of a secondary LSP are not used until the primary LSP fails. The resources allocated for a secondary LSP may be used by other LSPs until the primary LSP fails over to the secondary LSP. At that point, any set of LSPs that are using the resources for the secondary LSP must be preempted.

## Administrative Status Information

GMPLS introduces a new object/TLV for administrative status information. The use of this information is optional. The information can be used in two ways:

- To indicate the LSP's administrative state, such as “Administratively down,” “testing,” or “deletion in progress.” The nodes can use this information to allow local decisions, such as making sure an alarm is not sent if the LSP is put in a test mode. In RSVP-TE, this object is carried in the PATH and RESV messages.
- To send a request to set the LSP's administrative state. This request is always sent to the ingress nodes that act on the request. In RSVP-TE, this object is carried in a Notify message (discussed later, in the section “Notify Messages”).

## Separation of Control and Data Channels

In optical networks, the control and data channels need to be separated for multiple reasons, including these:

- Multiple links can be bundled.
- Some data channels cannot carry control information.
- The integrity of a data channel does not affect the integrity of control channels.

The following two sections discuss two critical issues for the separation of data and control channels.

### Interface Identification

In MPLS, a one-to-one association exists between the data and control channels (except for MPLS link bundling). In GMPLS, where such association does not exist, it is necessary to convey additional information in signaling to identify the particular data channel being controlled. GMPLS supports explicit data channel identification by providing interface identification information. GMPLS allows the use of several interface identification schemes, including IPv4 or IPv6 addresses, interface indexes, and component interfaces (established via configuration or a protocol such as LMP). In all cases, the choice of the data interface is indicated by the addresses and identifiers used by the upstream node.

### Fault Handling

Two new faults must be handled when the control channel is independent of the data channel:

- **Control channel fault**—A link or other type of failure that limits the ability of neighboring nodes to pass control messages. In this situation, neighboring nodes are unable to exchange control messages for a period of time. Once communication is restored, the underlying

signaling protocol must indicate that the nodes have maintained their state through the failure. The signaling protocol must also ensure that any state changes that were instantiated during the failure are synchronized between the nodes.

- **Nodal fault**—A node’s control plane fails and then restarts and loses most of its state information but does not lose its data forwarding state. In this case, both upstream and downstream nodes must synchronize their state information with the restarted node. For any resynchronization to occur, the node undergoing the restart needs to preserve some information, such as its mappings of incoming labels to outgoing labels.

## Notify Messages

GMPLS provides a mechanism to inform nonadjacent nodes of LSP-related failures using *Notify messages*. In optical networks, failure notification sometimes has to traverse transparent nodes to notify the nodes responsible for restoring failed connections (transparent nodes do not originate or terminate connections). This mechanism enables target nodes to be notified directly and more quickly of a network failure. The Notify message has been added to RSVP-TE. The Notify message includes the IP address of the node that needs to be notified. Other nodes in the path just pass on the message until it reaches the targeted node. The Notify message differs from the error messages Path-Error and Reservation-Error in that it can be “targeted” to a node other than the immediate upstream and downstream neighbor.

Another application of the Notify message is to notify when the control plane has failed while the data plane is still functional. GMPLS uses this mechanism to identify *degraded* links.

## Inclusion of Technology-Specific Parameters

The previous sections discussed the enhancements to signaling that allow GMPLS to control the different types of packet and nonpacket networks. GMPLS also allows the inclusion of technology-specific parameters that are carried in the signaling protocol in traffic parameter-specific objects. This section looks at how this applies to SONET/SDH. A description of parameters that are specific to optical transport network (OTN) technology is not included in this book.

The SONET/SDH traffic parameters specify a set for SONET (ANSI T1.105) and a set for SDH (ITU-T G.707), such as concatenation and transparency. Other capabilities can be defined and standardized as well. These traffic parameters must be used when SONET/SDH is specified in the LSP Encoding Type field of a generalized label request, discussed earlier in the section “The Generalized Label.” The SONET/SDH traffic parameters are carried in the SENDER\_TSPEC and FLOWSPEC objects of RSVP-TE and in SONET/SDH TLVs in CR-LDP.

Figure 8-11 shows how the SONET/SDH traffic parameters are organized. The Signal Type indicates the type of the elementary signal of the request LSP. Several parameters can be applied on the signal to build the final requested signals. These parameters are applied using the Request Contiguous Concatenation (RCC), Number of Contiguous Components (NCC), and Transparency fields included in the traffic parameter.

JUNIPER Exhibit 1003  
App. 6, pg. 205



**Figure 8-11** SONET/SDH Traffic Parameters

Signal Type (8 Bits)	RRC (8 Bits)	NCC (8 Bits)
NVC (16 Bits)		Multiplier (16 Bits)
Transparency (32 Bits)		

Examples of signal types for SONET/SDH include VT1.5, VT2, VT3, VT6, STS1, VC-11, VC-12, VC2, VC-3, and VC-4, plus other possible types, depending on the level of concatenation and transparency.

The RRC field, the NCC field, and the Number of Virtual Components (NVC) field are used to negotiate the type of concatenation and the number of signals that are to be concatenated. As mentioned in Chapter 2, “Metro Technologies,” concatenation can be applied to signals to form larger signals. Different types of concatenation, such as contiguous or virtual, can be applied, and the information is related in the signaling protocol.

**NOTE**

Transparency, in the context of SDH/SONET signals, refers to the overhead signals, such as the section overhead (SOH) and the line overhead (LOH) in the case of SONET. Transparency indicates which of these overhead fields needs to remain untouched when delivered to the other end of the LSP.

## Link Management Protocol

Future networks may consist of optical switches, data switches that are managed by GMPLS. Thousands of fibers may connect a pair of nodes, and hundreds of wavelengths may exist on each fiber. Multiple fibers and wavelengths can be bundled to form TE links. These links need a control channel to manage routing, signaling, and link connectivity and management. LMP is a link-control protocol that runs between neighboring nodes to manage TE links.

LMP was created to address the issues of link provisioning and fault isolation to improve and scale network manageability. With GMPLS, the control channel between two adjacent nodes is no longer required to use the same physical medium as the data channels between those nodes. A control channel can run on a separate IP management network, a separate fiber, or a separate wavelength. LMP allows for the decoupling of the control channel from the component links. As such, the health of the control channel does not necessarily correlate to the health of the data links, and vice versa.

LMP is designed to provide four basic functions to a node pair:

- **Control channel management**—A core function of LMP that is used to establish and maintain control channel connectivity between neighboring nodes. This consists of lightweight Hello messages that act as a fast keepalive mechanism between the nodes.
- **Link connectivity verification**—An optional LMP function that is used to verify physical connectivity of the data-bearing channels between the nodes and to exchange the interface IDs that are used in GMPLS signaling. The error-prone manual cabling procedures make LMP link connectivity verification very useful.
- **Link property correlation**—A core function of LMP that is designed to aggregate multiple ports or component links into a TE link and to synchronize the properties of the TE link. Link properties, such as link IDs for local and remote nodes, the protection mechanism, and priority, can be exchanged via LMP using the LinkSummary message between adjacent nodes.
- **Fault management and isolation**—An optional LMP function that provides a mechanism to isolate link and channel failures in both opaque and transparent networks, irrespective of the data format. Opaque nodes are nodes where channels can be terminated for the purpose of examining the headers and data. Transparent nodes are nodes where channels pass through without termination.

LMP requires that a pair of nodes have at least one active bidirectional control channel between them. This control channel may be implemented using two unidirectional control channels that are coupled using the LMP Hello messages. LMP allows backup control channels to be defined, such as using the data-bearing channels as backup in case of failure in the primary control channels.

## GMPLS Protection and Restoration Mechanisms

GMPLS introduces the necessary features in routing, signaling, and link management to support the fault management required in optical and electronic networks. Fault management requires the following capabilities:

- **Fault detection**—For optical networks, fault detection can be handled via mechanisms such as loss of light (LOL) and optical signal-to-noise ratio (OSNR) at the optical level, and bit error rate (BER), SONET/SDH Alarm Indicator Signal (AIS), or LOL at the SONET/SDH level.
- **Fault isolation**—For GMPLS, LMP can be used for fault isolation. The LMP fault-management procedure is based on sending ChannelActive and ChannelFail messages over the control channel. The ChannelActive message is used to indicate that one or more data-bearing channels are now carrying user data. The ChannelFail message is used to indicate that one or more active data channels or an entire TE link has failed.
- **Fault notification**—GMPLS uses the RSVP-TE Notify message to notify nodes of any possible failures. The Notify message can be used over the data-bearing links to indicate

JUNIPER exhibit 1003

App. 6, pg. 207

a failure in the control plane, or over the control channels to indicate a failure in the data plane. The notify request object can be carried in the RSVP PATH or RESV messages and indicates the IP address of the node that should be notified when generating an error message.

GMPLS uses the following protection mechanisms:

- **1+1 protection**—The data is transmitted simultaneously over the two disjoint paths. The receiver selects the working path based on the best signal.
- **1:1 protection**—A dedicated backup path is preallocated to protect the primary path.
- **M:N protection**—M backup paths are preallocated to protect N primary paths. However, data is not replicated onto a backup path, but only transmitted in case of failure on the primary path.

For 1:1 and M:N protection, the backup paths may be used by other LSPs. For 1+1 protection, the backup paths may not be used by other LSPs because the data is transmitted on both paths.

- **Span protection**—Intermediate nodes initiate the recovery that requires switching to an alternative path. As part of the GMPLS routing extensions, the link protection type is advertised so that span protection can be used.
- **Span restoration**—Intermediate nodes initiate the recovery that requires switching to an alternative path. The alternative path is dynamically computed.
- **Path protection**—End nodes initiate the recovery that requires switching to an alternative path. The end nodes switch to the backup path.
- **Path restoration**—End nodes initiate the recovery that requires switching to an alternative path. The backup path is dynamically calculated upon failure.

## Summary of Differences Between MPLS and GMPLS

As you've learned in this chapter, GMPLS extends MPLS to support non-packet/cell interfaces. The support of the additional TDM, lambda, and fiber interfaces impacts the basic LSP properties, such as how labels are requested and communicated and the unidirectional LSP behavior, error propagation, and so on.

Table 8-1 summarizes the basic differences between MPLS and GMPLS described in this chapter.

**Table 8-1** *Differences Between MPLS and GMPLS*

MPLS	GMPLS
Supports packet/cell-based interfaces only.	Supports packet/cell, TDM, lambda, and fiber.
LSPs start and end on packet/cell LSRs.	LSPs start and end on “similar type” LSRs (that is, PSC, L2SC, TDM, LSC, FSC).

*continues*

JUNIPER Exhibit 1003

App. 6, pg. 208

**Table 8-1** *Differences Between MPLS and GMPLS (Continued)*

<b>MPLS</b>	<b>GMPLS</b>
Bandwidth allocation can be done in any number of units.	Bandwidth allocation can only be done in discrete units for some switching capabilities such as TDM, LSC, and FSC.
Typical large number of labels.	Fewer labels are allocated when applied to bundled links.
No restrictions on label use by upstream nodes.	An ingress or upstream node may restrict the labels that may be used by an LSP along a single hop or the whole path. This is used, for example, to restrict the number of wavelengths that can be used in the case where optical equipment provides a small number of wavelengths.
Only one label format.	Use of a specific label on a specific interface. Label formats depend on the specific interface used, such as PSC, L2SC, TDM, LSC, FSC.
Labels are used for data forwarding and are carried within the traffic.	Labels are a control plane construct only in GMPLS and are not part of the traffic.
No need for technology-specific parameters, because this is applied to packet/cell interfaces only.	Supports the inclusion of technology-specific parameters in signaling.
Data and control channels follow the same path.	Separation of control and data channels
MPLS fast-reroute.	RSVP-specific mechanism for rapid failover (Notify message)
Unidirectional LSPs.	Bidirectional LSPs enable the following: <ul style="list-style-type: none"> <li>• Possible resource contention when allocating reciprocal LSPs via separate signaling sessions</li> <li>• Simplified failure restoration procedures</li> <li>• Lower setup latency</li> <li>• Lower number of messages required during setup</li> </ul>
Labels cannot be suggested by upstream node.	Allow a label to be suggested by an upstream node and can be overwritten by a downstream node (to prevent delays with setting optical mirrors, for example)

## Conclusion

As you have seen in this chapter, many extensions for routing, signaling, technology-specific parameters, and LMP allow the use of MPLS over non-packet/cell networks. Mechanisms such as link bundling and shared link groups are added to routing to influence the traffic trajectory and to take advantage of how the physical network topology is laid out. Signaling mechanisms such as the enhancements to the label allow the GMPLS label to be used as a control construct that indicates to the TDM/optical devices what circuits to switch and how to switch them. The introduction of LMP helps in the easy provisioning and protection of optical circuits by allowing channel link connectivity verification and fault management and isolation.



This appendix discusses the following topics:

- SONET/SDH Frame Formats
- SONET/SDH Architecture
- SONET/SDH Concatenation

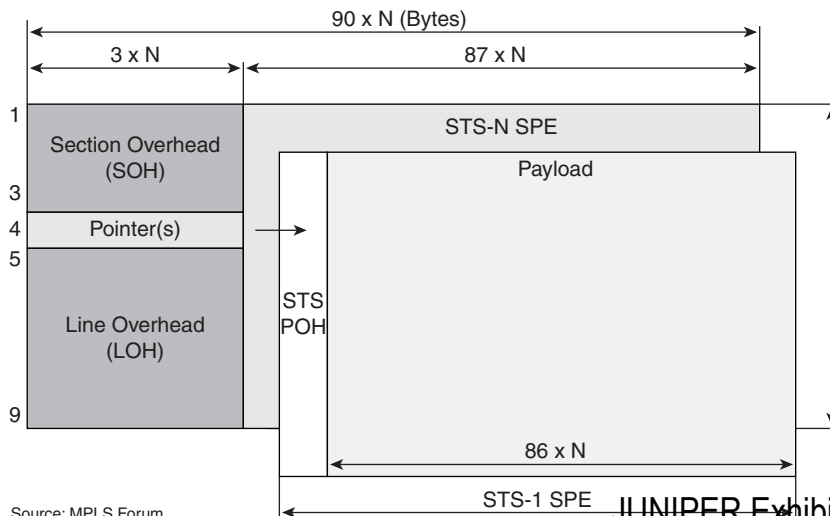
# SONET/SDH Basic Framing and Concatenation

This appendix explains basic SONET/SDH framing and concatenation. With the emergence of L2 metro services, SONET/SDH metro networks are being challenged to offer cost-effective and bandwidth-efficient solutions for transporting data services. The following sections describe the different elements of a SONET/SDH frame and how the elements can be combined to form bigger SONET/SDH pipes.

## SONET/SDH Frame Formats

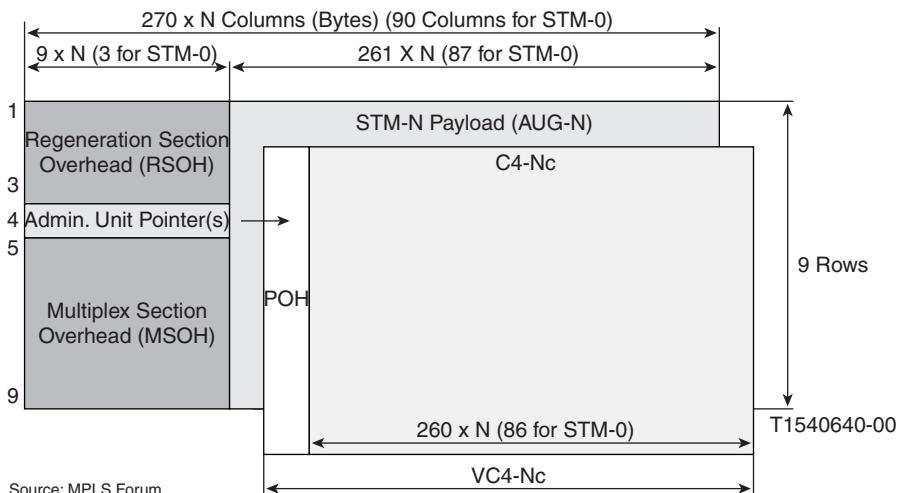
The fundamental signal in SONET is the STS-1, which operates at a rate of about 51 Mbps. The fundamental signal for SDH is STM-1, which operates at a rate of about 155 Mbps (three times the STS-1 rate). The signals are made of contiguous frames that consist of two parts: the transport overhead (TOH) contained in the header, and the payload. For synchronization purposes, the data can be allowed to shift inside the payload inside a Synchronous Payload Envelope (SPE) for SONET and inside the Virtual Container for SDH. The SPE inside the payload is referenced using a pointer. Figures A-1 and A-2 show the SONET and SDH frames.

**Figure A-1** SONET Frame Format



Source: MPLS Forum

Figure A-2 SDH Frame Format



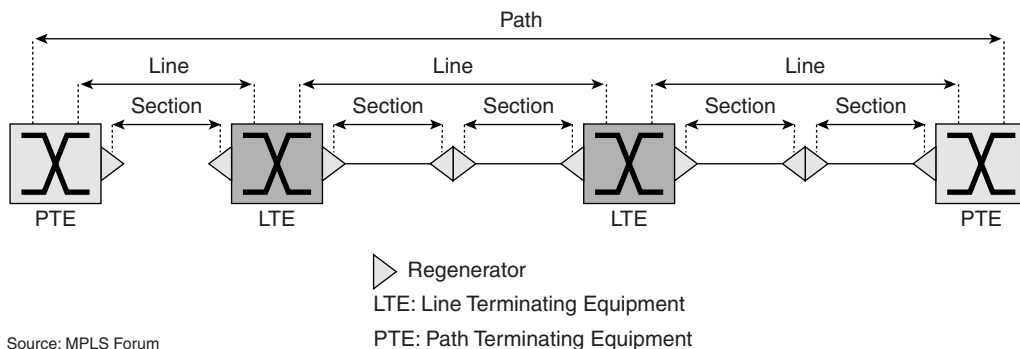
## SONET/SDH Architecture

The SONET/SDH architecture identifies three different layers, each of which corresponds to one level of communication between SONET/SDH equipment. The layers are as follows, starting with the lowest:

- The regenerator section, or section layer
- The multiplex section, or line layer
- The path layer

Figure A-3 shows the three SONET/SDH layers.

Figure A-3 SONET and SDH Layers

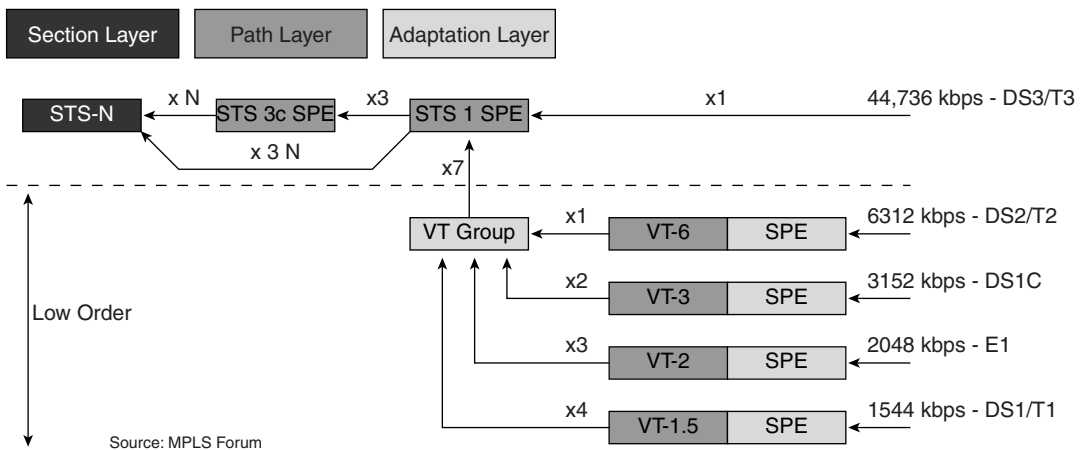


As shown in Figures A-1 and A-2, each of these layers in turn has its own overhead (header). The transport overhead (TOH) of a SONET/SDH frame is mainly subdivided into two parts



that contain the section overhead (SOH) and the line overhead (LOH). In addition, a pointer indicates the beginning of the SPE/Virtual Container in the payload of the overall frame. The SPE/Virtual Container itself is made up of the path overhead (POH) and a payload. This payload can be further subdivided into subelements, or a multiplex structure (signals). This multiplex structure leads to identifying time slots that contain tributary signals such as T1 (1.5 Mbps), E1 (2 Mbps), and so on. For example, a SONET STS-1 can be further divided into 7 \* VT-6 (virtual tributaries), where VT-6 is equal to 6.321 Mbps. A VT-6 can be divided into 4 \* VT 1.5, where a VT-1.5 is 1.544 Mbps or a T1. Figure A-4 shows the SONET multiplexing structure. Figures A-5 and A-6 show the SDH multiplexing structure. Table A-1 shows some helpful mapping between SONET and SDH.

**Figure A-4** SONET High-Order and Low-Order Multiplexing Structure



**Figure A-5** SDH High-Order Multiplexing Structure

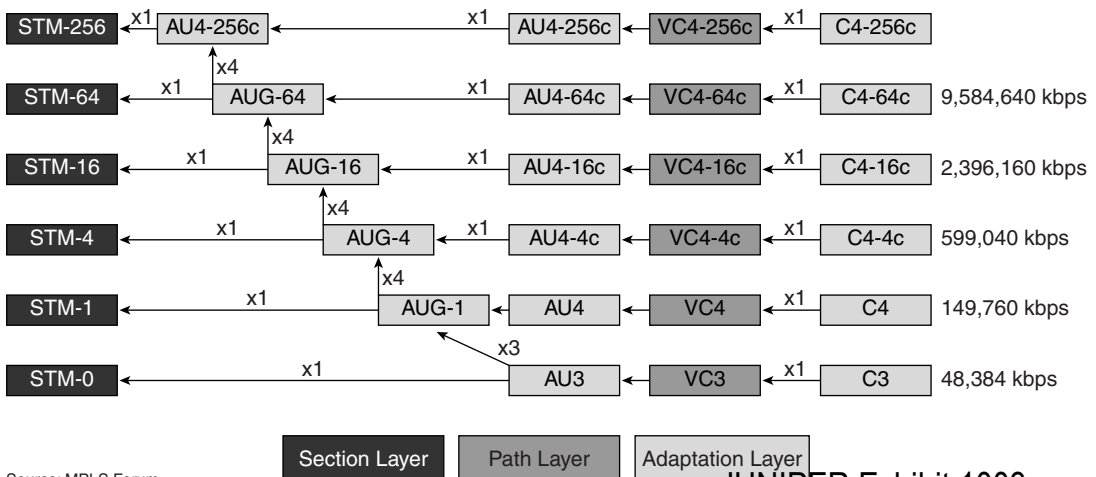
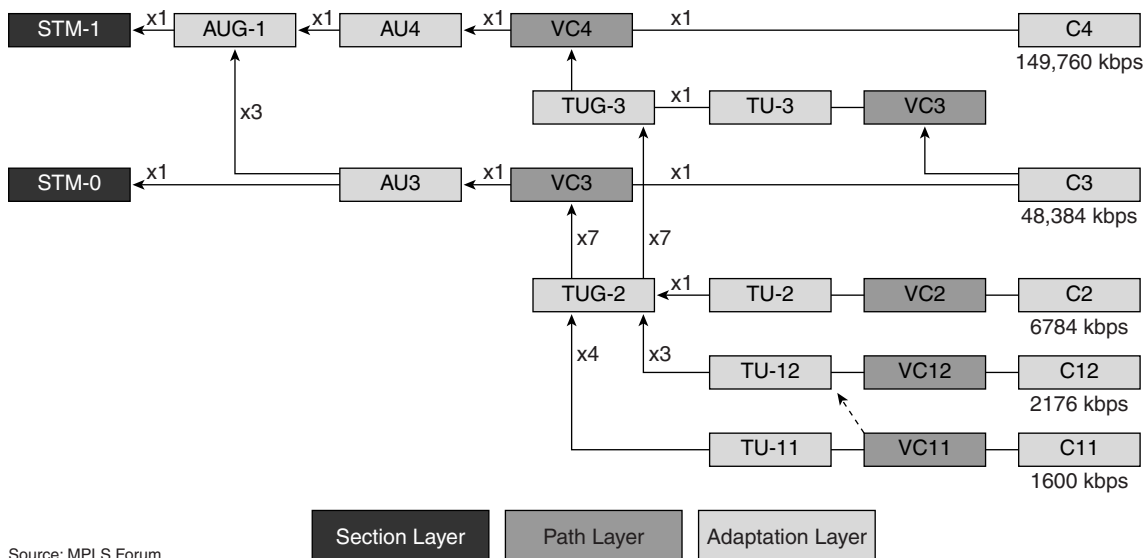


Figure A-6 SDH Low-Order Multiplexing Structure



Source: MPLS Forum

Table A-1 Helpful SONET/SDH Equivalency

SONET	SDH	
STS-1	VC-3	STM-0
STS-3c	VC-4	STM-1
VT-6	VC-2	
VT-3		
VT-2	VC-12	
VT-1.5	VC-11	
STS-12c	VC-4-4c	STM-4
STS-48c	VC-4-16c	STM-16
STS-192c	VC-4-64c	STM-64
STS-768c	VC-4-256c	STM-256

---

**NOTE** Note that STS-3, -12, -48, -192, -768, and so on are referred to as OC-3, -12, -48, -192, and so on.

---

An STS- $N$ /STM- $N$  signal is formed from  $N$  STS-1/STM-1 signals via byte interleaving. The SPEs/Virtual Containers in the  $N$  interleaved frames are independent and float according to their own clocking. This means that an STS-3 (OC3) pipe with bandwidth of about 155 Mbps is formed from three STS-1 signals. An STS-12 (OC12) pipe with bandwidth of about 622 Mbps is formed from 12 STS-1 signals. The STS-1 signals are independent.

## SONET/SDH Concatenation

To transport tributary signals in excess of the basic STS-1/STM-1 signal rates, the SPEs/Virtual Containers can be concatenated—that is, glued together. In this case, their relationship with respect to each other is fixed in time, and they act as one bonded pipe.

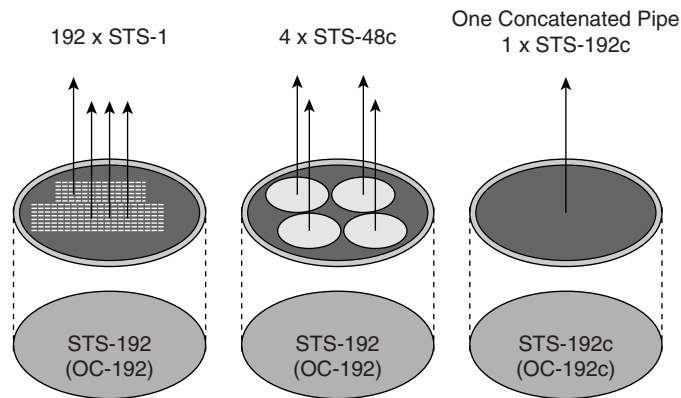
Different types of concatenations are defined, including *contiguous standard concatenation* and *virtual concatenation*.

### Contiguous Standard Concatenation

Contiguous standard SONET concatenation allows the concatenation of  $M$  STS-1 signals within an STS- $N$  signal, with  $M \leq N$  and  $M = 3, 12, 48, 192, 768$ , and so on in multiples of 4. The SPEs of these  $M$  STS-1s can be concatenated to form an STS- $Mc$ . The STS- $Mc$  notation is shorthand for describing an STS- $M$  signal whose SPEs have been concatenated ( $c$  stands for concatenated). This means that an STS-12c (OC12c) is formed from the concatenation of 12 STS-1 signals, and the 12 STS-1s act as one bonded pipe. Constraints are imposed on the size of STS- $Mc$  signals (that is, they must be a multiple of 3) and on their starting location and interleaving.

Figure A-7 shows an example of a SONET OC192 pipe (9.6 Gbps) that is multiplexed into 192 STS-1s or into four concatenated STS-48c pipes.

One of the disadvantages of standard concatenation is the lack of flexibility in starting time slots for STS- $Mc$  signals and in their interleaving. This means that the provider has to deploy SONET/SDH circuits with the predefined concatenation bandwidth size and with bandwidth increments that do not match its customer needs. This leads to inefficiencies in bandwidth deployment. Virtual concatenation solves this problem.

**Figure A-7** Sample SONET Structure

## Virtual Concatenation

Virtual concatenation is a SONET/SDH end-system service approved by the committee T1 of ANSI and ITU-T. The essence of this service is to have SONET/SDH end systems “glue” together the Virtual Containers or SPEs of separate signals rather than requiring that the signals be carried through the network as a single unit. In one example of virtual concatenation, two end systems that support this feature could essentially combine two STS-1s into a virtual STS-2c for the efficient transport of 100-Mbps Ethernet traffic. If instead these two end systems were to use standard concatenation with increments of STS-1, STS-3, and STS-12, a 100-Mbps pipe would not fit into an STS-1 (51 Mbps) circuit and would have to use an STS-3c (155 Mbps) circuit, therefore wasting about 55 Mbps of bandwidth. By using a virtual-concatenated STS-2c circuit (around 100 Mbps), the operator can achieve 100 percent efficiency in transporting a 100-Mbps Ethernet pipe.

### NOTE

The industry has suggested the use of *arbitrary contiguous concatenation*, which is similar in nature to virtual concatenation; however, it is applied inside the SONET/SDH network rather than the SONET/SDH end systems. Virtual concatenation will emerge as the solution of choice for next-generation data over SONET/SDH network deployments.

## Conclusion

This appendix has presented the basics of SONET/SDH framing and explained how the SONET/SDH technology is being adapted via the use of standard and virtual concatenation to meet the challenging needs of emerging data over SONET/SDH networks in the metro. The emergence of L2 metro services will challenge the legacy SONET/SDH network deployments and will drive the emergence of multiservice provisioning platforms (MSPPs) that will efficiently transport Ethernet, Frame Relay, ATM, and other data services over SONET/SDH.



JUNIPER Exhibit 1003  
App. 6, pg. 219

From the Library of Tal Lavian

# GLOSSARY

## A

---

**add/drop multiplexer (ADM).** A device installed at an intermediate point on a transmission line that enables new signals to come in and existing signals to go out. Add/drop multiplexing can be done with optical or electronic signals. The device may deal only with wavelengths, or it may convert between wavelengths and electronic TDM signals.

**adjacency.** A relationship formed between selected neighboring routers and end nodes for the purpose of exchanging routing information.

## B

---

**black hole.** Routing term for an area of the internetwork where packets enter but do not emerge due to adverse conditions or poor system configuration within a portion of the network.

**Building Local Exchange Carriers (BLECs).** Service providers that offer broadband services to businesses and tenants concentrated in building offices.

## C

---

**class of service (CoS).** A classification whereby different data packets that belong to a certain class receive similar quality of service.

**committed burst size (CBS).** A parameter associated with CIR that indicates the size up to which subscriber traffic is allowed to burst in profile and not be discarded or shaped.

**committed information rate (CIR).** The minimum guaranteed throughput that the network must deliver for the service under normal operating conditions.

**component link.** A subset of a bigger link. A channel within a SONET/SDH channelized interface is an example of a component link.

**control plane.** A logical plane where protocol packets get exchanged for the purpose of achieving multiple functions, such as setting up paths used for packet forwarding or for managing the nodes in the network.

**customer edge (CE) device.** A device such as a switch or router that resides at the customer premises. The device could be owned by the customer or the provider.

**customer premises equipment (CPE).** Terminating equipment, such as switches, routers, terminals, telephones, and modems, supplied by the telephone company, installed at customer sites, and connected to the telephone company network.

## D

---

**Data Packet Transport (DPT).** A Media Access Control protocol that adds resiliency and protection to packet networks deployed in a ring topology.

**Decoupled Transparent LAN Service (DTLS).** A service that emulates a LAN over an IP/MPLS network, similar to VPLS. DTLS, however, proposes to remove any L2 switching from the provider edge devices and restrict the L2 switching to the customer edge devices.

**detour LSP.** An LSP that is set up to reroute the traffic in case the main LSP fails.

**Diffserv.** A method used to classify IP packets so that different classes receive different quality of service treatment when forwarded in the network.

## E

---

**Ethernet LAN Service (E-LAN).** A multipoint-to-multipoint Ethernet service.

**Ethernet over MPLS (EoMPLS).** An L2 tunneling technique that allows Ethernet frames to be carried over an IP/MPLS network.

**Ethernet over SONET/SDH (EOS).** A technology that allows Ethernet packets to be transported over a SONET/SDH TDM network.

**Ethernet Virtual Connection (EVC).** A point-to-point Ethernet service.

**Explicit Route Object (ERO).** A field that indicates the path to be taken when traffic is forwarded.

## F

---

**fiber-switch capable (FSC) interfaces.** Interfaces that can forward data based on a position of the data in the real-world physical spaces. This is typical of optical cross-connects that switch traffic on the fiber or multiple-fiber level.

**Fixed Filter (FF).** Reservation style that creates a distinct reservation for traffic from each sender. This style is common for applications in which traffic from each sender is likely to be concurrent and independent. The total amount of reserved bandwidth on a link for sessions using FF is the sum of the reservations for the individual senders.



**Forwarding Information Base (FIB).** A data structure and way of managing forwarding in which destinations and incoming labels are associated with outgoing interfaces and labels.

**frame check sequence (FCS).** Extra characters added to a frame for error control purposes. Used in HDLC, Frame Relay, and other data link layer protocols.

## G

---

**generalized label request.** An MPLS label scheme that extends the use of the MPLS label to nonpacket networks.

**Generalized Multiprotocol Label Switching (GMPLS).** A generalized MPLS control plane that allows the provisioning and protection of circuits over both packet and nonpacket networks.

**Generic Attribute Registration Protocol (GARP).** A protocol defined by the IEEE to constrain multicast traffic in bridged Ethernet networks.

**Gigabit Ethernet (GE).** Standard for a high-speed Ethernet, approved by the IEEE 802.3z standards committee in 1996.

## I

---

**incumbent local exchange carrier (ILEC).** Traditional telephony company.

**interexchange carrier (IXC).** Common carrier that provides long-distance connectivity between dialing areas serviced by a single local telephone company.

**interface.** In routing or transport terminology, a network connection or a port.

**Interior Gateway Protocol (IGP).** Internet protocol used to exchange routing information within an autonomous system. Examples of common Internet IGPs include IGRP, OSPF, and RIP.

## L

---

**L2TPv3.** An L2 tunneling protocol that allows the tunneling of Ethernet packets over an L3 IP network.

**label block.** A block of MPLS labels exchanged between two MPLS routers.

**Label Distribution Protocol (LDP).** A standard protocol between MPLS-enabled routers to negotiate the labels (addresses) used to forward packets. The Cisco proprietary version of this protocol is the Tag Distribution Protocol (TDP).

**label switch router (LSR).** Forwards packets in an MPLS network by looking only at the fixed-length label.

**label switched path (LSP).** A path that MPLS packets traverse between two edge LSRs.

**lambda-switch capable (LSC) interfaces.** Interfaces that can forward data based on the wavelength on which it was received. This is typical of optical cross-connects that switch traffic on the wavelength level.

**Layer 2-switch capable (L2SC) interfaces.** Interfaces that can recognize L2 cell or frame boundaries and can forward data based on L2 headers. This is typical of interfaces on ATM switches, Frame Relay switches, and L2 Ethernet switches.

**Link Aggregation Control Protocol (LACP).** A protocol that allows multiple Ethernet links to be bundled in a larger pipe.

**link bundling.** Aggregating multiple links into a bigger pipe.

**Link Management Protocol (LMP).** Establishes and maintains control channel connectivity between neighbors. LMP also enables neighbor discovery, which allows neighbors to identify connected devices, obtain UNI connectivity information, and identify and verify port-level connections, network-level addresses, and corresponding operational states for every link.

**LSP tunnel.** A configured connection between two routers that uses MPLS to carry the packets.

## M

---

**maximum transmission unit (MTU).** Maximum packet size, in bytes, that a particular interface can handle.

**Media Access Control (MAC) address.** Standardized data link layer address that is required for every port or device that connects to a LAN. Other devices in the network use these addresses to locate specific ports in the network and to create and update routing tables and data structures. MAC addresses are 6 bytes long and are controlled by the IEEE.

**metropolitan area network (MAN).** Network that spans a defined metropolitan or regional area; smaller than a WAN but larger than a LAN.

**multidwelling units (MDUs).** Buildings that contain multiple housing units, such as apartment complexes and university dormitories.

**multiple service operator (MSO).** Cable service provider that also provides other services, such as data and voice telephony.

**multiplexing.** Scheme that allows multiple logical signals to be transmitted simultaneously across a single physical channel.

**multipoint-to-multipoint (MP2MP).** An any-to-any connection between end systems.

**Multiprotocol Label Switching (MPLS).** Switching method that forwards IP traffic using a label. This label instructs the routers and switches in the network where to forward the packets based on pre-established IP routing information.

**multitenant units (MTUs).** Multitenant building offices that are recipients of broadband services by a BLEC.

## N

---

**Network Management System (NMS).** System responsible for managing at least part of a network. An NMS is generally a reasonably powerful and well-equipped computer, such as an engineering workstation. NMSs communicate with agents to help keep track of network statistics and resources.

**Network-to-Network Interface (NNI).** A specification of the interface between a backbone system and another backbone system. For example, the specification of an optical interface that connects two optical switches in the carrier network.

**Notify message.** A message used by RSVP-TE to notify other nodes of certain failures.

## O

---

**OAM&P.** Operations, administration, maintenance, and provisioning. Provides the facilities and personnel required to manage a network.

**optical cross-connect (OXC).** A network device that switches high-speed optical signals.

## P

---

**packet multiplexing.** Data packets coming in from different locations and being multiplexed over the same output wire.

**packet-switch capable (PSC).** Systems such as IP/MPLS routers that can switch data packets.

**Packet-switch capable (PSC) interfaces.** Interfaces that can recognize packet boundaries and can forward data based on packet headers. This is typical of interfaces on routers and Layer 3 Ethernet switches.

**packet switching.** The ability to forward packets in the network based on packet headers or fixed labels.

**peak information rate (PIR).** Specifies the maximum rate above the CIR at which traffic is allowed into the network and may get delivered if the network is not congested.

**point of local repair (PLR).** The router at which a failed LSP can be locally rerouted.

**point-to-point (P2P).** A one-to-one connection between two end systems.

**provider (P) device.** Normally, a core IP/MPLS router that offers a second level of aggregation for the provider edge devices.

**provider edge (PE) device.** A provider-owned device that offers the first level of aggregation for the different customer edge (CE) devices.

**pseudowire (PW).** A representation of packet-leased line, or a virtual circuit between two nodes.

## Q

---

**Q-in-Q.** An Ethernet encapsulation technique that allows Ethernet packets that already have an 802.1Q VLAN tag to be 802.1Q VLAN tagged again.

## R

---

**Record Route Object (RRO).** A field that indicates the path that traffic takes when forwarded.

**regional Bell operating company (RBOC).** Regional telephone company formed by the breakup of AT&T.

**Resilient Packet Ring (RPR).** A Media Access Control standard protocol that adds resiliency and protection to packet networks deployed in a ring topology.

**Resource Reservation Protocol (RSVP).** Protocol that supports the reservation of resources across an IP network. Applications running on IP end systems can use RSVP to indicate to other nodes the nature (bandwidth, jitter, maximum burst, and so on) of the packet streams they want to receive.

**RSVP-TE.** A protocol that extends RSVP to support traffic engineering over an IP/MPLS network.

## S

---

**Shared Explicit (SE).** Reservation style that allows a receiver to explicitly select a reservation for a group of senders, rather than one reservation per sender, such as in the FF style. Only a single reservation is shared between all senders listed in the particular group.

**shared risk link group (SRLG).** A grouping that indicates similar risk characteristics for a set of elements. A set of fibers, for example, that share the same conduit belong to the same SRLG, because if the conduit is cut, all fibers will fail.

**shortest path first (SPF) algorithm.** Routing algorithm that iterates on length of path to determine a shortest-path spanning tree. Commonly used in link-state routing algorithms. Sometimes called Dijkstra's algorithm.

**SONET/SDH terminal multiplexer (TM).** A device installed at an endpoint on a transmission line that multiplexes multiple transmission lines, such as DS1s/DS3s, into a SONET/SDH network.

**spanning tree.** Loop-free subset of a network topology.

**Synchronous Digital Hierarchy (SDH).** A standard for delivering data over optical fiber. SDH is used in Europe.

**Synchronous Optical Network (SONET).** A standard for delivering data over optical fiber. SONET is used in North America and parts of Asia.

**Synchronous Payload Envelope (SPE).** The payload-carrying portion of the STS signal in SONET. The SPE is used to transport a tributary signal across the synchronous network. In most cases, this signal is assembled at the point of entry to the synchronous network and is disassembled at the point of exit from the synchronous network. Within the synchronous network, the SPE is passed on intact between network elements on its route through the network.

## T

---

**time-division multiplexing (TDM).** Technique in which information from multiple channels can be allocated bandwidth on a single wire based on preassigned time slots. Bandwidth is allocated to each channel regardless of whether the station has data to transmit.

**time-division multiplexing (TDM) interfaces.** Interfaces that can recognize time slots and can forward data based on the data's time slot in a repeating cycle. This is typical of interfaces on digital cross-connects, SONET ADMs, and SONET cross-connects.

**Time To Live (TTL).** A mechanism to prevent loops in IP networks. The TTL field gets decremented every time a packet traverses a router. When TTL reaches 0, the packet can no longer be forwarded.

**traffic engineered (TE) link.** A link that is set up to divert the traffic over a path different than what is calculated by Interior Gateway Protocols.

**traffic engineering (TE).** Techniques and processes that cause routed traffic to travel through the network on a path other than the one that would have been chosen if standard routing methods were used.

**traffic engineering (TE) tunnel.** A label-switched tunnel that is used for traffic engineering. Such a tunnel is set up through means other than normal L3 routing; it is used to direct traffic over a path different from the one that L3 routing could cause the tunnel to take.

**traffic trunk.** Physical and logical connection between two switches across which network traffic travels. A backbone is composed of a number of trunks.

**Transparent LAN Service (TLS).** A service that extends the LAN over the MAN and WAN.

**trunk.** Physical and logical connection between two switches across which network traffic travels. A backbone is composed of a number of trunks.

**tunnel.** A connection between two end systems that allows the encapsulation of packets within it.

## U

---

**unidirectional path switched ring (UPSR).** Path-switched SONET rings that employ redundant, fiber-optic transmission facilities in a pair configuration. One fiber transmits in one direction, and the backup fiber transmits in the other. If the primary ring fails, the backup takes over.

**unnumbered link.** A link that does not have an IP address assigned to it.

**User-to-Network Interface (UNI).** A specification of the interface between an end system and a backbone system. An example is the specification of an Ethernet interface that connects a switch at the customer site and a router at the provider site.

## V

---

**virtual circuit (VC).** Logical circuit created to ensure reliable communication between two network devices. A virtual circuit is defined by a VPI/VCI pair, and can be either permanent (PVC) or switched (SVC).

**Virtual Container.** An SDH signal that transports payloads that are smaller than an STM-0 (48,384 kbps) payload. VC is part of the SDH hierarchy.

**virtual LAN (VLAN).** Group of devices on one or more LANs that are configured (using management software) so that they can communicate as if they were attached to the same wire, when in fact they are located on a number of different LAN segments. Because VLANs are based on logical instead of physical connections, they are extremely flexible.

**Virtual Private LAN Service (VPLS).** A service that extends the notion of a switched Ethernet LAN over an IP/MPLS network.

**virtual router forwarding (VRF).** A VPN routing/forwarding instance. A VRF consists of an IP routing table, a derived forwarding table, a set of interfaces that use the forwarding table, and a set of rules and routing protocols that determine what goes into the forwarding table. In general, a VRF includes the routing information that defines a customer VPN site that is attached to a PE router.

**Virtual Tributary.** A SONET signal that transports payloads that are smaller than an STS-1 (44,736 kbps) payload. VT is part of the SONET hierarchy.

## W

---

**waveband.** A set of contiguous wavelengths that can be switched together as a unit.

**wavelength-division multiplexing (WDM).** Optical technology whereby multiple optical wavelengths can share the same transmission fiber. The spectrum occupied by each channel must be adequately separated from the others.

**Wildcard Filter (WF).** Reservation style in which a single shared reservation is used for all senders to a session. The total reservation on a link remains the same regardless of the number of senders.



JUNIPER Exhibit 1003  
App. 6, pg. 229

From the Library of Tal Lavian



## A

---

access 17  
 to building risers 12  
 control 38  
 EOS with packet multiplexing 30  
 Ethernet 20

add operation 36

add/drop multiplexers (ADMs) 6

addresses  
 MAC 46–47, 95  
 switching 45–50  
 VPN-IPv4 79

adjacency, FA 170

ADM, EOS function inside 24

administration  
 fault 188  
 LMP 187  
 networks 153  
 provisioning 153  
 NFS 154

administrative status information 185

ADMs (add/drop multiplexers) 6

aggregation, trunks into tunnels 130

AIS (Alarm Indicator Signal) 188

Alarm Indicator Signal (AIS) 188

algorithms, path computation 174

all-Ethernet networks, building 14

all-to-one bundling 61

arbitrary contiguous concatenation, 208. *See also*  
 concatenation

architecture  
 GMPLS 151–152  
   interfaces 167–168  
   need for 152–157  
 SONET/SDH 204–207

Asian markets 16

attributes  
 MEF 50–68  
 routes 78  
 site of origin 79  
 traffic  
   parameters 128  
   trunks 127

augmented model 159. *See also* signaling

autodiscovery 103

availability 54  
 avoiding loops 94

## B

---

backbones  
 L3VPNs 74–81  
 scaling 69  
 traffic forwarding 80

backdoor loops 95

backward trunks 127

bandwidth  
 encoding 183  
 Ethernet 9  
 requirements 27  
 RPR 35–39  
 VCAT 25–28

benefits of link bundling 173

BER (bit error rate) 188

BGP (Border Gateway Protocol) 104

bidirectional LSPs 183

billing 153

binding labels 137–138

bit error rate (BER) 188

BLECs (Building Local Exchange  
 Carriers) 9

block offset 109

breadth of services 13

breaking loops 94

broadcast 47

broadcast frames 57

broadcast storms 41

building all-Ethernet data networks 14

Building Local Exchange Carriers (BLECs) 9

building risers, access to 12

bundling links 172

Bundling service attribute 61

bypass tunnels 149

## C

---

cable modem termination systems (CMTSs) 35

capabilities, interface switching descriptors 175–177

carriers, metro Ethernet

CBS (Committed Burst Size) 53  
 CCAMP (Common Control and Measurement Plane) 151  
 CE (customer edge) 73  
 central office (CO) 97  
 centralized provisioning, TDM networks 153–154  
 centralized switching, EOS with 32  
 channels
 

- control separation 185
- data separation 185
- fault handling 185–186

 CIR (Committed Information Rate) 53  
 circuits, provisioning 153  
 class of service (CoS) parameters 55  
 CMTSs (cable modem termination systems) 35  
 CO (central office) 97  
 Committed Burst Size (CBS) 53  
 Committed Information Rate (CIR) 53  
 Common Control and Measurement Plane (CCAMP) 151  
 comparisons
 

- Frame Relay 20
- GMPLS/MPLS 189–190

 components
 

- links, 172
- VPN 73–74

 computation of paths 156, 174  
 concatenation
 

- SONET/SDH 207–208
- VCAT 25–28

 configuration
 

- hub-and-spoke (Gigabit Ethernet) 42
- L2PE 113
- PE 113
- PW 83

 congestion, TE 122  
 connections
 

- EVC 50
- LCAS 28
- last-mile 20
- metro. *See* metro
- multitpoint 20
- VPN components 73–74

 constrained-based routing LDP (CR-LDP) 134  
 constraint-based routing 129–130  
 contiguous standard SONET concatenation, 207. *See also* concatenation, SONET/SDH  
 control channels
 

- fault handling 185

separating 185  
 CoS (class of service) parameters 55  
 costs, overbuilding networks 12  
 CPE (customer premises equipment) 7  
 CR-LDP (constrained-based routing LDP) 134  
 cross-connect (XC) 27  
 customer edge (CE) 73  
 customer premises equipment (CPE) 7  
 customers, restrictions to number of 69

---

## D

data channels, separating 185  
 data equipment, EOS interfaces in 34  
 Data Packet Transport (DPT) 35  
 data-link connection identifier (DLCI) 142  
 Decoupled Transparent LAN Service (DTLS) 111  
 delay 54  
 delivering L3VPNs over IP 74–81  
 deployment
 

- Ethernet L2 services 82
- incumbents 13–15
- international 15–16
- legacy 70, 71
- metro 5–8
- RPR 36
- services 35–39

 descriptors, interface switching capability 175–177  
 detour LSPs 148  
 devices
 

- CE 73
- non-bridging as spokes 101
- P 74
- PE 73

 DiffServ codepoints (DSCPs) 56  
 Digital Wrapper (DW) 181  
 discovery, resource topologies 156  
 DLCI (data-link connection identifier) 142  
 documentation. *See* MEF  
 DPT (Data Packet Transport) 35  
 drop operation 36  
 DSCPs (Diffserv codepoints) 56  
 DTLS (Decoupled Transparent LAN Service) 111  
 dual-homed MTU devices 102  
 DW (Digital Wrapper) 181  
 dynamic provisioning model 154–157

**E**

early service providers of metro Ethernet 9–13  
 eBGP (External BGP) 103  
 edge, core 17  
 E-LAN (Ethernet LAN Service) 51  
 ELS (Ethernet Line Service) 51  
 emulation, links 83  
 encapsulation
 

- Ethernet 86
- GRE 74
- VPLS 93

 encoding
 

- bandwidth 183
- LSP types 181

 end-to-end repair method 147  
 Enterprise Systems Connection (ESCON) 24  
 EOS (extended operating system) 24
 

- interfaces 34
- packet multiplexing at access 30
- packet switching 31–33
- transport services 28–30

 ERO (EXPLICIT\_ROUTE) object 137, 142  
 ESCON (Enterprise Systems Connection) 24  
 establishing trunks, RSVP-TE 134–146  
 Ethernet 8–9
 

- access 20
- early service providers of 9–13
- Gigabit Ethernet. *See* Gigabit Ethernet
- L2VPN services 19
- MEF 50–68
- services
  - over IP/MPLS networks 81–83
  - over MPLS 85–90
  - PW 83–85
  - VPLS 90–116
- over SONET/SDH 23–25
- transport 39–42

 Ethernet LAN Service (E-LAN) 51  
 Ethernet Line Service (ELS) 51  
 Ethernet physical interface attribute 52  
 Ethernet Virtual Connection (EVC) 50  
 European markets 16  
 EVC (Ethernet Virtual Connection) 50  
 existing legacy TDM infrastructure 14  
 expanding capacity 122  
 explicit label control 184

explicit paths, establishing 135  
 EXPLICIT\_ROUTE object (ERO) 137, 142  
 extensions, RSVP 135  
 External BGP (eBGP) 103  
 extranets 75

**F**

FA (forwarding adjacency) 170  
 facility backups 149  
 fairness, RPR 38  
 fast provisioning 9  
 fast reroute, MPLS 146–149  
 fault handling 185–186  
 fault management 188  
 FCS (frame check sequence) 85  
 FF (Fixed Filter) 139  
 FIB (Forwarding Information Base) 97  
 fiber connectivity (FICON) 24  
 fibers, SRLG 175  
 fiber-switch capable (FSC) interfaces 168  
 FICON (fiber connectivity) 24  
 fields, TLV 184  
 filters 61, 139  
 first mile 5  
 Fixed Filter (FF) 139  
 flexibility of service 11  
 flooding 47  
 FLOW\_SPEC object 145  
 formatting
 

- frames (SONET/SDH) 203
- loop-free topologies 93

 forward operation 36  
 forward trunks 127  
 forwarding
 

- packets 77
- paths 135
- tables 76
- traffic 80

 forwarding adjacency (FA) 170  
 Forwarding Information Base (FIB) 97  
 frame check sequence (FCS) 85  
 Frame Relay
 

- comparisons 20
- VPNs 105

- frames 87
  - service frame delivery attribute 56
  - SONET/SDH 203
- FSC (fiber-switch capable) interfaces 168
- full mesh
  - loop avoidance 94
  - LSPs 97
- functions
  - EOS 24
  - NSP 85
  - Q-in-Q 59

## G

- GARP (Generic Attribute Registration Protocol) 57
- generalized labels 180, 181
- Generalized MPLS. *See* GMPLS
- Generalized Payload Identifier (G-PID) 181
- Generic Attribute Registration Protocol (GARP) 57
- Generic Framing Protocol (GFP) 90
- generic path selection 128
- generic routing encapsulation (GRE) tunnels 74
- geography, variations of deployment 5
- GFP (Generic Framing Protocol) 90
- Gigabit Ethernet
  - hub-and-spoke configuration 42
  - rings 40
- global access control 38
- GMPLS (Generalized MPLS) 151–152
  - inclusion of technology-specific 186–187
  - interfaces 167–168
  - LMP 187
  - MPLS 189–190
  - need for 152–157
  - protection 188–189
  - restoration 188–189
- G-PID (Generalized Payload Identifier) 181
- granularity of bandwidth 9
- GRE (generic routing encapsulation) tunnels 74
- greenfield value proposition 11–13
- groups, SRLG 175

## H-J

- hierarchies
  - LSP 177–180
  - LSPs 170
  - VPLS 97
- hops
  - loose 143
  - strict 143
- hub-and-spoke configuration (Gigabit Ethernet) 42
- iBGP (Internal BGP) 103
- identification
  - interfaces 185
  - PW 84
- IGP (Internet Gateway Protocol) metrics 123
- ILEC (incumbent local exchange carrier) 6
- implementation of RSVP 135
- inclusion of technology-specific parameters 186–187
- incumbent local exchange carrier (ILEC) 6
- incumbents, deployment of 13–15
- infrastructure
  - signaling models 157–159
  - SONET/SDH (Ethernet over) 23–25
- infrastructure existing legacy TDM 14
- installation of physical ports 153
- interfaces
  - EOS 34
  - Frame Relay 20
  - FSC 168
  - GMPLS 167–168
  - identification 185
  - L2 20
  - L2SC 168
  - LSC 168
  - NNI 157
  - PSC 167
  - switching 175, 176, 177
  - TDM 168
  - UNI 50, 157
- Internal BGP (iBGP) 103
- international deployment 15–16
- Internet Gateway Protocol (IGP) metrics 123
- interworking with legacy deployments 70–71
- intranets 75

IP/MPLS networks  
 Ethernet services over 81–83  
 MPLS 85–90  
 PW 83–85  
 VPLS 90–116  
 IPv4, VPN-IPv4 addresses 79  
 IS-IS Traffic Engineering (IS-IS-TE) 169  
 jitter 55

## L

L2 (Layer 2)  
 backbones 69  
 configuring 113  
 interfaces 20  
 labels 116  
 L2SC (Layer 2 switch capable) interfaces 168  
 L2TP control connection endpoints (LCCEs) 84  
 L2VPN BGP model 106–107  
 L2VPN services 19  
 L3VPN services 74–81  
 Label Distribution Protocol (LDP) 86  
 label switched path. *See* LSP  
 label switched routers (LSRs) 135  
 LABEL\_REQUEST object 137, 141  
 labels  
 binding 137–138  
 explicit control 184  
 generalized 180–181  
 L2PE 116  
 optimizing 180  
 ranges 109  
 sets 182  
 suggested 183  
 switching  
 GMPLS 167–168  
 nonpackets 159  
 troubleshooting 184  
 WAN 115  
 LACP (Link Aggregation Control Protocol) 57  
 lambdas, SRLG 175  
 lambda-switch capable (LSC) interfaces 168  
 LAN (local-area network)  
 resources 18  
 VPLS 90–116  
 large enterprises (LEs) 5  
 last mile 5, 20  
 Layer 2 (L2)  
 backbones 69  
 configuring 113  
 interfaces 20  
 labels 116  
 Layer 2 switch capable (L2SC) interfaces 168  
 Layer Control Processing packets 57  
 layers, SONET/SDH 204–207  
 LCAS (Link Capacity Adjustment Scheme) 28  
 LCCEs (L2TP control connection endpoints) 84  
 LDP (Label Distribution Protocol) 86  
 BGP signaling 104  
 CR-LDP 134  
 directly connected PEs 87  
 learning  
 MAC addresses 46–47, 95  
 qualified/unqualified 97  
 legacy deployments, interworking with 70–71  
 LEs (large enterprises) 5  
 line overhead (LOH) 187  
 Link Aggregation Control Protocol (LACP) 57  
 Link Capacity Adjustment Scheme (LCAS) 28  
 Link Management Protocol (LMP) 187  
 links  
 bundling 172  
 emulating 83  
 GMPLS protection/restoration 188–189  
 protection types 174  
 unnumbered 171  
 LMP (Link Management Protocol) 187  
 local access control 38  
 local-area network. *See* LAN  
 local link identifiers 172  
 local switching, EOS with 32, 33  
 logical separation, PE 77  
 LOH (line overhead) 187  
 LOL (loss of light) 188  
 loop-free topologies, creating 93  
 loops  
 avoiding 94  
 backdoor 95  
 loose hops 143

- loss 55
- loss of light (LOL) 188
- LSP (label switched path) 80, 126, 135
  - bidirectional 183
  - encoding types 181
  - full mesh 97
  - hierarchies 170, 177–180
  - RSVP tunnels 136
  - traffic trunks
- LSRs (label switched routers) 135

## M

---

- MAC (Media Access Control)
  - addresses 95
  - learning 46–47
  - switching 45–50
  - RPR 35
- management
  - fault 188
  - LMP 187
  - networks 153
  - NFS 154
- Maximum Burst Size (MBS) 53
- Maximum Transmit Unit. *See* MTU
- MBS (Maximum Burst Size) 53
- MDUs (multidwelling units) 5
- Media Access Control. *See* MAC
- MEF (Metro Ethernet Forum) 50–63
  - example of 63–68
- messages
  - Notify 186
  - PATH 141–145
  - RESV 145
- metrics, IGP 123
- metro
  - data view of 16–17
  - deployment 5–8
  - Ethernet 8–13
  - MEF 50–68
  - services 17–19
- Metro Ethernet Forum. *See* MEF
- migration, EOS as transport services 28–30
- modification
  - routing 168–186
  - signaling 168–186
- monitoring services 69
- MP2MP (multipoint-to-multipoint) 82
- MP-BGP (multiprotocol BGP) 77
- MPLS (Multiprotocol Label Switching)
  - Ethernet over 85–90
  - fast reroute 146–149
  - GMPLS 189–190
  - TE 125–130
- MPLS L3VPN 80
- MPLS L3VPNs 75
- MSOs (multiple service operators) 35
- MTU (Maximum Transmit Unit) 87, 99–100
- MTUs (multitenant units) 5
- multicast 47
- multicast frames 56
- multidwelling units (MDUs) 5
- multiple fibers, SRLG 175
- multiple lambdas, SRLG 175
- multiple LSPs, nesting 177–180
- multiple service operators (MSOs) 35
- multiple services 26
- multiple switching capabilities 175–177
- multiple T1 (nXT1) 7
- multiple TE links, bundling 172
- multiplexing 35, 61
- multiplexing packets at access 30
- multipoint connectivity 20
- multipoint-to-multipoint (MP2MP) 82
- multiprotocol BGP (MP-BGP) 77
- multitenant units (MTUs) 5

## N

---

- Native Service Processing (NSP) 85
- need for GMPLS 152–157
- nesting multiple LSPs 177–180
- Network Layer Reachability Information (NLRI) 109
- Network Management System (NMS) 154
- networks
  - all-Ethernet 14
  - GMPLS 151–157

## IP/MPLS

- Ethernet services over 81–83
- over MPLS 85–90
- over VPLS 90–116
- PW 83–85

management 153

metro. *See* metro

on-net 6

OTN 186

overbuilding 12

reliability

- equal-cost multipath (TE) 124
- IGP metrics (TE) 123
- routing (TE) 124
- TE 121–122
- techniques (TE) 123
- VC overlays (TE) 124

ring (RPR) 35–39

signaling models 157–159

SONET/SDH 183

TDM

- label switching 160–163
- provisioning 153–154

troubleshooting 68–71

trunking 48

WAN labels 115

WDM 163

NLRI (Network Layer Reachability Information) 109

NMS (Network Management System) 154

NNI (Network-to-Network Interface) 157

nodal fault handling 186

node pairs, LMP 187

non-bridging devices as spokes 101

non-directly connected PEs 88

nonpacket label switching 159

non-PSC networking devices 152, 157. *See also*

routers; TDM

Notify messages 186

NSP (Native Service Processing) 85

numbers, unnumbered links 171

nXT1 (multiple T1) 7

## O

O4 189

objects

ERO 142

FLOW\_SPEC 145

LABEL\_REQUEST 137, 141

RRO 144

SENDER\_TEMPLATE 145

SESSION 145

SESSION\_ATTRIBUTE 144

OC12 (12 STS-1s) 26

one-to-one backup method 148

on-net networks 6

Open Shortest Path First for Traffic Engineering  
(OSPF-TE) 169

operations

packets 36

traffic trunks 127

optical cross-connects (OXC) 152

optical networks

GMPLS 151–157

signaling models 157–159

optical signal-to-noise ratio (OSNR) 188

optical transport network (OTN) 186

optimization

bandwidth (RPR) 35–39

labels 180. *See also* labels

routing 168. *See also* routing

signaling 177. *See also* signaling

TE 121–122

equal-cost multipath 124

IGP metrics 123

routing 124

techniques 123

VC overlays 124

OSNR (optical signal-to-noise ratio) 188

OSPF-TE (Open Shortest Path First for Traffic  
Engineering) 169

OTN (optical transport network) 186

overbuilding networks 12

overlays

models 158. *See also* signaling

VC 124

overprovisioning 122

OXC (optical cross-connect) 152

## P

- P (provider) 74
- P2P (point-to-point) 82
- packets
  - DPT 35
  - EOS
    - switching 31–33
    - as transport services 28–30
  - forwarding 77
  - Layer Control Processing 57
  - multiplexing at access 30
  - nonpacket label switching 159
  - operations 36
  - PW 83
  - RPR 35–39
  - switching 45–50
- packet-switch capable (PSC) interfaces 157, 167
- packet-switched network (PSN) tunnels 80
- parameters
  - GMPLD 186–187
  - MEF 50–68
  - traffic attributes 128
- PATH message 141–145
- path overhead (POH) 204
- paths
  - computation 156, 174
  - establishing 135
  - generic selection 128
  - LSP 80
  - LSPs 135
  - RSVP-TE 134–146
- pay as you grow model 11
- PE (provider edge) 73
  - configuring 113
  - LDP with directly connected 87
  - logical separation 77
  - non-directly connected 88
- Peak Information Rate (PIR) 53
- peer model 159. *See also* signaling
- penultimate hop popping 90
- performance. *See also* optimization
  - parameters 54
  - TE 121–122
    - IGP metrics 123
    - techniques 123
- PE-rs 101
- physical ports, installing 153
- PIR (Peak Information Rate) 53
- planning provisioning 153
- POH (path overhead) 204
- point-to-point (P2P) 82
- policing attribute 129
- POP (point of presence) 8
- ports
  - physical 153
  - trunks 48
- preemption attribute 128
- pricing
  - models 13
  - overbuilding networks 12
  - services 15
- priority attribute 128
- protection
  - GMPLS 188–189
  - link types 174
  - TLV fields 184
- protocols
  - BGP signaling 104
  - GFP 90
  - GMPLS 151–152, 157
  - IP 74–81
  - MAC (RPR) 35
  - modifying 168–186
  - RSVP
    - extensions 135
    - implementing 135
  - SRP 38
  - STP 50, 94
- provider (P) 74
- provider edge (PE) 73
  - configuring 113
  - LDP with directly connected 87
  - logical separation 77
  - non-directly connected 88
- provisioning
  - dynamic 154–157
  - path computation 156
  - TDM networks 153–154
- provisioning services 69
- PSC (packet-switch capable) interfaces 157, 167
- PSN (packet-switched network) tunnels 80
- PW (pseudowire) 83–85



## Q-R

---

Q-in-Q function 59  
 qualified learning 97

ranges, labels 109  
 raw mode 86  
 RD (Route Distinguisher) 79, 103  
 reach of services 13  
 RECORD\_ROUTE object (RRO) 138, 144  
 reducing TDM bandwidth 25–28  
 reference models, VPLS 90  
 regulations 5, 15  
 reliability 68–71  
 reliability of TE 121–122
 

- equal-cost multipath 124
- IGP metrics 123
- routing 124
- techniques 123
- VC overlays 124

 remote link identifiers 172  
 reordering 87  
 requirements
 

- bandwidth 27
- VPLS 91

 resilience attribute 128  
 resiliency, RPR 36  
 resilient packet ring (RPR) 35–39  
 resource attributes 129  
 resource discovery topologies 156  
 resource reservation 135, 138
 

- FF 139
- SE 140
- WF 140

 resources, LAN 18  
 restoration, GMPLS 188–189  
 restrictions to number of customers 69  
 RESV message 145  
 retail models 10  
 rings
 

- Gigabit Ethernet 40
- RPR 35–39
- steering 37
- wrapping 37

 risk, SRLG 175  
 Route Distinguisher (RD) 79, 103  
 route reflectors 105

route target (RT) 103  
 routers
 

- attributes 78
- interfaces, GMPLS 167–168
- RSVP-TE 134–146

 routing
 

- constraint-based 129–130
- CR-LDP 134
- fast reroute 146–149
- GMPLS protection/restoration 188–189
- GRE 74
- LSP 80, 170
- modifying 168–186
- TE 124

 RPR (resilient packet ring) 35–39  
 RRO (RECORD\_ROUTE) object 138, 144  
 RSVP (Resource Reservation Protocol)
 

- extensions 135
- implementing 135
- LSP tunnels 136

 RSVP-TE 134–146  
 RT (route target) 103

## S

---

scalability 9, 68–71
 

- backbones 69
- VPLS 97

 SDH (Synchronous Digital Hierarchy)
 

- architecture 204–207
- concatenation 207–208
- frame formats 203

 SDL (Simple Data Link) 24  
 SE (Shared Explicit) 139–140  
 section overhead (SOH) 187, 204  
 security 61  
 SENDER\_TEMPLATE object 145  
 separation of control/data channels 185  
 service frame delivery attribute 56  
 service multiplexing attribute 61  
 service providers, metro Ethernet 9–13  
 services 17–19
 

- data view of 16–17
- Ethernet
  - over IP/MPLS networks 81–83
  - over MPLS 85–90

JUNIPER Exhibit 1003

App. 6, pg. 238

- PW 83–85
  - VPLS 90–116
  - flexibility of 11
  - MEF 50–68
  - monitoring 69
  - multiple 26
  - pricing 15
  - provisioning 69
  - RPR 35–39
  - time to bring up 11
  - transport (EOS) 28–30
  - VPN components 73–74
  - SESSION object 145
  - SESSION\_ATTRIBUTE object 138, 144
  - sets, labels 182
  - Shared Explicit 139–140
  - shared risk link group (SRLG) 175
  - signaling
    - BGP 104
    - GMPLS protection/restoration 188–189
    - models 157–159
    - modifying 168–186
    - SONET/SDH concatenation 207–208
    - VPLS 93
  - Simple Data Link (SDL) 24
  - site of origin 79
  - small and medium-sized businesses (SMBs) 5
  - small office/home office (SOHO) 5
  - SMBs (small and medium-sized businesses) 5
  - SOH (section overhead) 187, 204
  - SOHO (small office/home office) 5
  - SONET
    - architecture 204–207
    - concatenation 207–208
    - frame formats 203
  - SONET/SDH
    - Ethernet over 23–25
    - label switching 162
    - LSP 161
    - LSRs 161
    - networks 183
  - Spanning Tree Protocol (STP) 50, 94
  - spans 5
  - SPE (Synchronous Payload Envelope) 23, 203
  - special reuse protocol (SRP) 38
  - speed, Frame Relay speed 20
  - spokes, non-bridging devices as 101
  - SRLG (shared risk link group) 175
  - SRP (special reuse protocol) 38
  - stacking 58
  - static provisioning, TDM networks 153–154
  - steering, ring 37
  - STP (Spanning Tree Protocol) 50, 94
  - strict hops 143
  - STS-1 (50 Mbps) 27
  - styles, resource reservation 138. *See also* resource reservation
  - subobjects 142, 144
  - suggested labels 183
  - support
    - VC 20
    - waveband switching 182
  - swapping 60
  - switches
    - EOS functions inside 24
    - GMPLS 151–152
      - need for 152–157
    - LMP 187
  - switching 45–60
    - interfaces 175–177
    - labels
      - GMPLS 167–168
      - nonpackets 159
    - LSP 80
    - packets (EOS) 31–33
    - types 181
    - waveband 182
  - Synchronous Payload Envelope (SPE) 23, 203
- 
- ## T
- 
- tables
    - forwarding 76
    - MAC learning 46–47
  - tagged mode 86
  - tagging VLAN 49–50, 57
  - targets
    - RT 103
    - VPNs 78
  - TDM (time-division multiplexing) 6
    - existing legacy infrastructure 14

- GMPLS 151–152
  - need for 152–157
- interfaces 168
- label switching 160–163
- network provisioning 153–154
- TE (traffic engineering) 121–122
  - equal-cost multipath 124
  - IGP metrics 123
  - IS-IS-TE 169
  - MPLS 125–130
  - OSPF-TE 169
  - routing 124
  - techniques 123
  - tunnels (MPLS fast reroute) 146–149
  - VC overlays 124
- time to bring up service 11
- time-division multiplexing. *See* TDM
- TLS (Transparent LAN Service) 81
- TLV (type length value) 97, 184
- TOH (transport overhead) 204
- topologies
  - loop-free 93
  - resource discovery 156
- traffic
  - forwarding 80
  - parameters 52, 128
  - trunks 135
    - attributes/operations 127
    - comparing to LSPs 126
- traffic engineering. *See* TE
- translation, VLAN 60
- Transparent LAN Service (TLS) 81
- transport, Ethernet 39–42
- transport overhead (TOH) 204
- transport services, EOS 28–30
- troubleshooting 80, 184
  - congestion 122
  - fault management 188
  - LCAS 28
  - MPLS L3VPN 80
  - networks 68–71
  - Notify messages 186
- trunks, 48
  - RSVP-TE 134–146
  - traffic
    - attributes/operations 127
    - comparing to LSPs 126

- tunnels
  - aggregation 130
  - bypass 149
  - GRE 74
  - LSP hierarchies 177–180
  - MPLS fast reroute 146–149
  - PSN 80
  - RSVP LSP 136
- type length value (TLV) 97, 184
- types
  - of link protection 174
  - of LSP encoding 181
  - of PW 84
  - of service providers 5
  - of VC 88
  - switching 181

## U

- U.S. incumbent landscape 13–15
- UNI (User-to-Network Interface) 50, 157
- unicast frames 56
- unnumbered links 171
- unqualified learning 97
- untagged Ethernet packets 49–50
- User-to-Network Interface (UNI) 50, 157

## V

- values, TLV fields 184
- variations of deployment 5
- VC (virtual circuit)
  - overlays 124
  - support 20
  - types 88
- VCAT (virtual concatenation) 25–28
- VCI (virtual connection identifier) 142
- virtual circuit. *See* VC
- virtual path identifier (VPI) 142
- Virtual Private LAN (VPLS) 90–116
- virtual router forwarding (VRF) 76
- VLAN (virtual LAN)
  - switching 45–50
  - tagging 49–50, 57
- VLAN Tag Preservation 58

- VLAN Tag Translation or Swapping 60
- VPI (virtual path identifier) 142
- VPLS (Virtual Private LAN) 90–116
- VPN (virtual private network)
  - components 73–74
  - Frame Relay 105
  - L2VPN services 19
  - L3VPNs over IP 74–81
  - of origin 79
  - target 78
- VPN-IPv4 addresses 79
- VRF (virtual router forwarding) 76

## W-X

---

- WAN (wid-area network) labels 115
- waveband switching 182
- WDM (wavelength division multiplexing) 35, 163
- WF (Wildcard Filter) 139, 140
- wholesale models 10
- wrapping, ring 37
  
- XC (cross-connect) 27

# The LSP Protection/Restoration Mechanism in GMPLS

by

Ziying Chen

The LSP Protection/Restoration Mechanism in GMPLS

by

Ziying Chen

A graduation project submitted to  
the Faculty of Graduate and Postdoctoral Studies  
in partial fulfillment of the requirements for the degree of  
Master of Computer Science

School of Information Technology and Engineering

University of Ottawa

Ottawa, Ontario, Canada, K1N 6N5

October 1, 2002

## **Abstract**

This report introduces the new switching technology Generalized Multiprotocol Label Switching (GMPLS) and traffic engineering. It outlines the components of the GMPLS path protection/restoration mechanism, and it specifies how different protocols contribute to path protection/restoration in GMPLS. This report specifies different path protection/restoration mechanisms. It illustrates how they work and how the signaling protocol supports them. Also, some case studies are provided to illustrate how the recovery mechanism is constructed in practice. At the end, the report compares these path protection/restoration mechanisms and introduces the current trend of protection/restoration in the industry.

## Acknowledgements

I would like to thank my supervisor, Professor Gregor Bochmann, for his support and care during my study under his supervision. His guidance is very appreciated!

I would like to thank my uncle, Chi Kan Leung. His long-term support makes my study dream in Canada come true.

I would like to thank my family for their encouragement and care.

I would like to thank my grandmother and all the relatives in our big family for their moral support and help.

I would also like to thank my friends that study with me throughout the years in the school.

And, I would also like to thank all the people at the University of Ottawa I have had the pleasure to meet.



## Table of Contents

1. Introduction.....	6
2. Overview of GMPLS.....	7
2.1 LSP Hierarchy.....	8
2.2 The Mesh Network.....	11
2.3 Traffic Engineering.....	12
2.4 The GMPLS Control Plane.....	13
2.4.1 Resource Discovery.....	14
2.4.2 Enhancements in the Routing Protocol to Support GMPLS.....	14
2.4.3 Enhancements in MPLS Signaling to Support GMPLS .....	20
2.4.4 Path Computation.....	22
3. Overview of Path Protection/Restoration.....	25
4. Multiple Protocols Contribute to GMPLS LSP Protection/Restoration.....	27
4.1 OSPF Extensions.....	27
4.1.1 Extensions to OSPF for supporting Traffic Engineering.....	28
4.1.2 Extensions to OSPF for supporting GMPLS.....	31
4.1.2.1 Unnumbered link support in OSPF.....	31
4.1.2.2 Shared Risk Link Group (SRLG) .....	32
4.1.2.3 Link Protection Type.....	32
4.1.2.4 Interface Switching Capability Descriptor.....	33
4.2 Link Management Protocol (LMP) .....	33
4.3 GMPLS Signaling .....	37
4.3.1 GMPLS signaling: RSVP-TE with extensions.....	37
4.3.1.1 Signaling Support for Fault Notification.....	47
4.3.2 GMPLS signaling: CR-LDP with extensions.....	48
4.4 The Hello Protocol.....	51
5. The Recovery Mechanism in GMPLS.....	52
5.1 Protection Mechanisms.....	52
5.1.1 Local Protection.....	55
5.1.1.1 Link Protection.....	55
5.1.1.2 Node Protection.....	57
5.1.2 Global Protection .....	57
5.2 Restoration Mechanisms.....	58
5.2.1 Local Restoration.....	58
5.2.2 Global Restoration.....	65
6. Case Studies.....	66
6.1 Case Study 1: The end-to-end LSP Protection.....	66
6.2 Case Study 2: The Domain-Specific Protection.....	72
6.3 Case Study 3: The Link-layer Protection and Local Reroute.....	78
7. Conclusion.....	81
References.....	84

## 1. Introduction

Multiprotocol Label Switching (MPLS) [1] is a recent switching technology that has been proposed for IP networks with two main objectives: (a) providing a more efficient mechanism for packet forwarding than traditional routing, and (b) providing tools for quality of service and traffic engineering. It is based on a switching principle very similar to ATM cell switching (VPI/VCI correspond to labels) and Time-Division Multiplexing (time slots correspond to labels).

With the increased traffic within the Internet, there is a tendency to have backbone connections with very high bandwidth capability, including optical fibers possibly with Dense Wavelength Division Multiplexing (DWDM). The principle of DWDM is again very similar to time-division multiplexing (wavelengths correspond to labels).

It can be foreseen that the future data networks will include various switching techniques at various levels of the capacity hierarchy, from optical transmission up to the packet level. Since the switching techniques expected to be used at these different levels, that is, MPLS, optical space switching, DWDM, and time division multiplexing, all require that a logical connection between the source and the destination must be established before the data can be sent, it has been proposed that it would be good if the same signaling protocol could be used for controlling the establishment of such logical connections at all these different levels. While the signaling protocols at these different levels may not be completely identical because they may require certain level-dependent parameters, nevertheless, the logical structure and most of the message content could be identical for the signaling at these different levels. General MPLS (GMPLS) [2] is intended as such a signaling protocol that could be used at these different switching levels.

With the development of networks, new technologies provide high bandwidth capacity, which makes a significant data loss if a failure cannot be recovered timely. It is imperative for GMPLS networks to provide protection/restoration of traffic.

This report gives an introduction to the general area and provides an overview of the GMPLS protocol and related standards. The main emphasis of this report is on the path protection/restoration mechanisms that can be used with GMPLS. It specifies how different protocols contribute to path protection/restoration in GMPLS, including signaling and routing protocols. This report specifies different path protection/restoration mechanisms. It illustrates how they work and how the signaling protocol supports them. It also addresses some problems remaining to be solved, and provides some answers. Some case studies are provided to illustrate how the recovery mechanism can be used in practice. At the end, this report compares these path protection/restoration mechanisms and introduces the current status of path protection/restoration mechanisms in the industry.

## 2. Overview of GMPLS

MPLS evolved from several similar technologies that were invented in the middle of the 1990s, for example, *IP switching* by Ipsilon [3] [4] [5], *Tag Switching* by Cisco [6], *Aggregated Route-based IP Switching* by IBM [7], and *Cell Switching Router* by Toshiba [8]. They all use label swapping to forward data, and they all use IP addressing and IP-based routing protocols like OSPF. At the end of the 1990s, the Internet Engineering Task Force (IETF) standardized the technology and named it MPLS [1].

A label is a short, fixed-length entity and it does not encode any information from the network layer header. A node that supports MPLS is called Label Switching Router (LSR). A label is inserted in front of each data packet on the entry in the network. At each LSR, the packet is forwarded based on the value of the label, and forwarded to an outgoing interface with a new label. In some situations, the incoming interface is also a factor to determine the outgoing interface. This operation is called Label Swapping. When the data packet arrives at the destination node, the label is stripped off and the packet is handed to the upper layer to process. The path that data is forwarded by label swapping across a network is called Label Switched Path (LSP). In the illustration in Figure 1.0, the LSP is (Node1, Node2, Node3). The head node of the LSP is called ingress node, e.g., Node 1, and the ending node of the LSP is called egress node, e.g., Node 3.

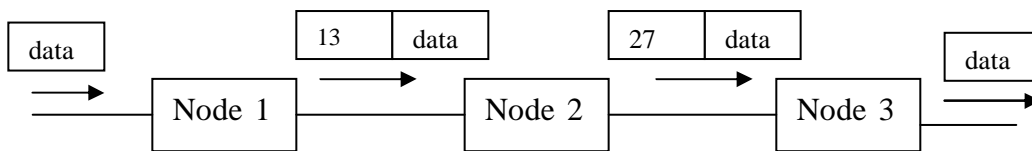


Figure 1.0: data is forwarded along the LSP

The function of forwarding can be partitioned into two components: control component and forwarding component. The forwarding component is responsible for the forwarding of data from the input port to the output port in a router according to the forwarding table. The control component is responsible for the construction and maintenance of the forwarding table. These two components are also named forwarding plane and control plane.

MPLS [1] provides routers with the label switching technology to forward data. The router can make a forwarding decision based on two sources of information: the label forwarding table and the label carried in the data. Based on the incoming label (and maybe also the incoming interface), the forwarding table provides enough information to forward the data, e.g., outgoing label, outgoing interface, and so on (see Figure 1.1 for a forwarding entry).

Incoming information	Outgoing information
Incoming label	Outgoing label
...	Outgoing interface
	...

Figure 1.1 the logical view of an entry in the forwarding table.

MPLS supports data forwarding based on a label. The original MPLS architecture [1] assumes that a Label Switching Router (LSR) has a forwarding plane which can (a) recognize packet (or cell) boundaries, and (b) process packet (or cell) headers. However, there are routers that cannot recognize packet boundaries or process packet headers, e.g., TDM switches, optical cross-connects (OXC), etc. But different label modeling techniques can allow these routers (switches) to forward data using the same principle of label switching. For example, the time slot of TDM, the lambda (or wavelength) of a WDM switch, the port of an OXC, etc, can be modeled as a label. That means the forwarding plane is different, but the control plane can be same. Such a technology is called Generalized MPLS (GMPLS) [2]. GMPLS extends MPLS. With GMPLS, a switch whose forwarding plane recognizes neither packet nor cell boundaries can also forward data using this extended label switching technology. GMPLS supports multiple types of switching: packet (cell), TDM, lambda, and space (port) switching. This means that GMPLS can forward data based on time slots, wavelengths, physical ports and labels.

GMPLS models wavelength, TDM channels or time slots as labels [9], and the name *generalized label* refers to all these different “labels” [10].

## 2.1 LSP Hierarchy

So far, GMPLS supports five types of interfaces (see [2]).

### (1) Packet Switch Capable (PSC) interfaces

They are interfaces that can recognize packet boundaries and can forward data based on the content of the packet header. An example is an Ethernet interface of an IP router, which can recognize the header boundary of an IP packet.

### (2) Layer-2 Switch Capable (L2SC) interfaces

They are interfaces that recognize frame/cell boundaries and can forward data based on the content of the frame/cell header. An example is an interface of an ATM switch that forwards cells based on the label encoded by ATM VCI/VPI.

(3) Time-Division Multiplex Capable (TDM) interfaces

They are interfaces that forward data based on the data's time slot in a repeating cycle. An example is an interface of a SONET switch.

(4) Lambda Switch Capable (LSC) interfaces

They are interfaces that forward data based on the wavelength on which the data is received. An example includes the interface of an Optical Cross-Connect (OXC), which can distinguish lambdas.

(5) Fiber-Switch Capable (FSC) interfaces

They are interfaces that forward data based on a position of the data in the real world physical spaces. An example is an interface of a Photonic Cross-Connect (PXC), which can operate on a per-fiber basis.

We can see that interfaces (3), (4) and (5) are unable to check the content of the user data, while (1) and (2) can process the packet (cell) headers.

A circuit can be established only between, or through, interfaces of the same type. Depending on the particular technology being used for each interface, different circuit names can be used, e.g. SONET/SDH circuit, light-path, etc. In the context of GMPLS, all these are referred to a common name: Label Switched Path (LSP).

In MPLS, LSPs can be nested, e.g., several LSPs of the same level can be multiplexed into a single LSP of another level. The nested LSP concept in MPLS has been extended to GMPLS [11]. A new LSP is multiplexed inside an existing higher-order LSP so that the preexisting LSP serves as a link along the path of the new LSP [12]. This is referred to as LSP hierarchy. The ordering of LSPs is based on the link multiplexing capabilities of the nodes. A hierarchical LSP can be established using the same type of interface, or between different types of interface.

A hierarchical LSP can be established if an interface is capable of multiplexing several LSPs from the same technology (layer). For example, 4 OC-48 links can be multiplexed into an OC-192 link. A lower order SDH/SONET LSP (OC-48) can be nested in a higher order SDH/SONET LSP (OC-192) (see Figure 1.2).

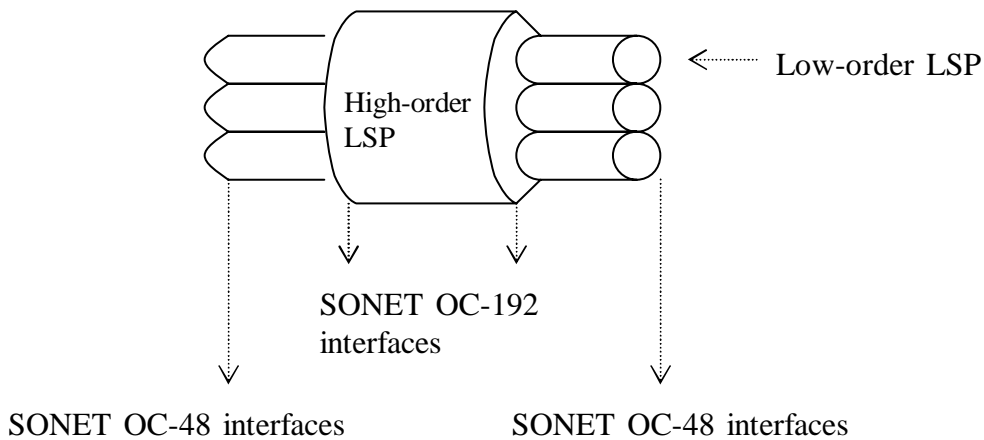


Figure 1.2: a hierarchical LSP is established on the same type of interfaces.

A hierarchical LSP can also be established between different types of interface. Let us discuss the following example. An LSP which starts and ends on Packet Switch Capable (PSC) interfaces can be nested (together with other LSPs) into an LSP which starts and ends on SONET (TDM) interfaces – assuming that the SONET interfaces have bigger capacity. That LSP which starts and ends on SONET interfaces can again be nested into an LSP which starts and ends on Lambda Switch Capable (LSC) interfaces.

Figure 1.3 shows an example where nested LSPs occur between different types of interfaces.

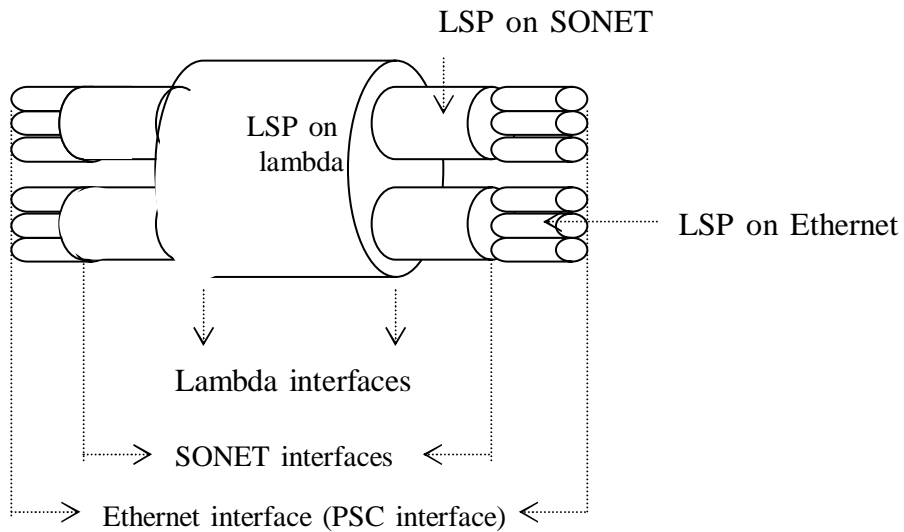


Figure 1.3: LSP hierarchy between different interfaces

At the top of this LSP hierarchy is the LSP with FSC interfaces, followed by LSC, then by TDM, L2SC and PSC interfaces (the reversed order of the above 5 interfaces). So, an

LSP which starts and ends on PSC interfaces can be nested into an LSP which starts and ends on L2SC interfaces. This LSP, further, can be nested into an LSP that starts and ends on TDM interfaces, which further can be nested into an LSP that starts and ends on LSC interfaces. Again, the LSP starts and ends on LSC interfaces can further be nested into an LSP that starts and ends on FSC interfaces. The example in Figure 1.3 shows a three-level hierarchical LSP. For each level of a given hierarchy, there is a separate control instance. The LSP is independently computed based on that level of routing information, and independently signaled. Examples follow in the subsequent sections.

## 2.2 The Mesh Network

The trend of the Internet transport infrastructure is to have an optical network core interconnecting high-speed routers (and switches) (see [13]).

A lightpath is a point-to-point optical layer connection between two access points in an optical network (see [14] for the definition). An example is shown in Figure 1.4. A wavelength connects two edge OXCs through two ports of the OXCs. Note that the two edge OXCs may be bridged by a number of OXCs and the wavelength may be switched by these transit OXCs. The lightpath is referred to as an LSP in the context of GMPLS if the lightpath is set up by GMPLS signaling.

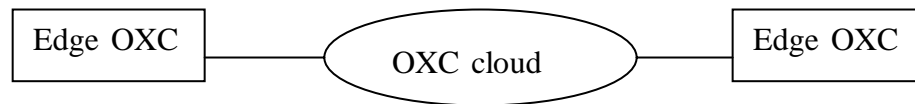


Figure 1.4: a lightpath

This report only considers the LSP recovery mechanism in a mesh network. An example of a mesh network is shown in Figure 1.5. In the example, LSRs which are packet-switch capable (called PSC LSRs) are connected to SONET switches. And the SONET switches are connected to an optical core network. One PSC LSR is connected to its peer over dynamically established LSPs across the optical core. The optical core is assumed to be incapable of processing packet headers. It is also assumed that a path must be established across the optical core network before the PSC LSRs can communicate.

The optical core network consists of OXCs that are connected by point-to-point optical links. The OXC can operate at the level of individual wavelength. The OXCs are mesh-connected (to form a general topology). Each node has the GMPLS-implemented control plane. What does it mean? (a) The nodes can forward data using label switching. For example, OXCs can forward data by label switching - based on the input wavelength, which is modeled as a label, to make a forwarding decision. (b) Each node uses GMPLS signaling (e.g., RSVP-TE with extensions) and GMPLS routing protocols (e.g. OSPF-TE with extensions).

It is recommended that the optical network control plane should utilize IP-based protocols (e.g., signaling and routing) for dynamic provisioning and restoration of light-paths within

and across optical networks. This is because signaling and routing mechanisms developed for IP traffic engineering applications can be reused in optical networks [15].

The OXC provides lambda-switch capable interfaces, and the multiplexing capacity of the interface is usually much bigger than that of the packet-switch capable interface. Furthermore, wavelength (or lambda) cannot multiplex packets directly. Therefore, SONET switches, e.g., OC-192/OC-48 switches, provide the optical core network access to the PSC LSRs. In this example and rest of this report, it is assumed that the edge OXC has interfaces that provide WDM capabilities for lambda-switch capable interfaces, also it has interfaces that provide SONET section level signals (e.g., OC-192 including all overheads). The SONET switch can multiplex a number of same-level LSPs that deliver packets into a single SONET path. The SONET switch also has a GMPLS-implemented control plane – it uses label switching to forward data and GMPLS signaling and routing protocols.

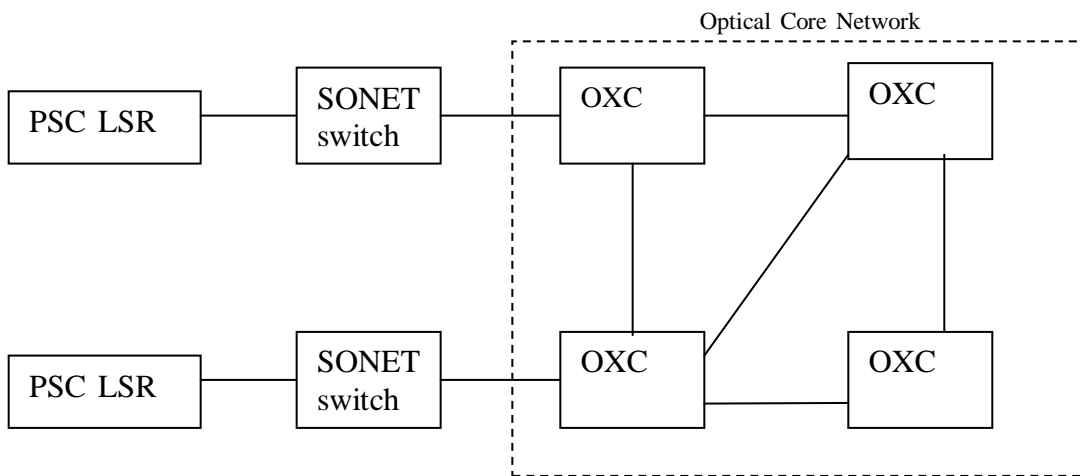


Figure 1.5: a GMPLS mesh network example

### 2.3 Traffic Engineering

The task of mapping traffic flows onto an existing physical network topology to optimize the network resource utilization and facilitate the network operations is called Traffic Engineering (see [16] for detailed definition). Traffic Engineering (TE) provides the ability to move traffic flow away from congestions and onto a potentially less congested physical path across a network.

TE properties are information used to support traffic engineering. For example, TE properties for a link include: available bandwidth, maximum bandwidth, etc.

Traditional routing protocols (e.g., OSPF) do not consider Traffic Engineering and they have been extended to advertise TE properties in a network by IETF, e.g., *TE LSAs to extend OSPF for Traffic Engineering* [17], *OSPF Extensions to Support Multi-Area Traffic Engineering* [18]. For example, assuming that two routers are connected by a link, with the TE information advertised by the extended OSPF, both routers understand



the available bandwidth of the link, the maximum bandwidth of the link, etc. Each router stores the TE properties in a database, which are learnt from the advertisement provided by the routing protocol. With the TE properties in the database, a node understands the TE properties of the network. And the database of the routers will be synchronized within the entire routing area. The information in the database can be used for a path computation algorithm to compute a path across the network to meet the Traffic Engineering requirements.

The Traffic Engineering (TE) link concept is introduced with the current development of traffic engineering and optical networks. A TE link is a logical link that has TE properties [19]. The Internet draft [19] explains the meaning of “logical”: it is a way to group/map the information about certain physical resources (and their properties) into the information that is used by CSPF for the purpose of path computation, and by GMPLS signaling. Both ends of the link must do the mapping/grouping consistently. By “consistent”, it means the information advertised by one end of the link does not conflict with that advertised by the other end of the link. Examples of a TE link are: a physical link, an LSP, or a bundle of physical links. The TE properties of a TE link are exchanged like traditional link information by routing protocols, e.g., carried by OSPF advertisement messages.

As we said, an LSP can be regarded as a TE link. Because of the benefits introduced by optical networks, e.g., high bandwidth, the capacity of an LSP constructed by lambdas likely cannot be utilized completely by one user. The routing protocol can advertise this LSP as a TE link into the routing domain, which can be used for the path computation algorithm to calculate paths, path aggregation (e.g., shared by other LSPs that require a portion of the LSP capacity), etc. We say that there is a “forwarding adjacency” (FA) between the end-nodes of the advertised LSP [20]. And such an LSP is named FA-LSP [20]. As a TE link, the TE properties are also associated with the FA-LSP.

In a hierarchical LSP, the high-order LSPs tunnel low-order LSPs. The high-order LSP should be advertised by the routing protocol as a TE link (and they become a FA-LSP), so that the unreserved bandwidth is utilized.

We will see examples of the TE link and FA-LSP in the subsequent sections.

## **2.4 The GMPLS Control Plane**

There are five major functions in the control plane of GMPLS: resource discovery, routing, path computation, link management and signaling. We briefly introduce these functions here and we will specify the portions of these functions that are related to this report in the subsequent sections.

Resource discovery is the procedure through which nodes within a network find out the resource in the network. It provides the information for signaling and path computation. Path computation uses an algorithm to calculate an explicit-routed LSP (ER-LSP).

The routing function uses the IP-based routing protocols to distribute and maintain the information about the topology and resources of the network. The routing protocol is the means by which non-local resources are discovered. The topology and resources of the network will be taken into account as parameters for the path algorithm to calculate an ER-LSP.

Signaling is the procedure through which service provisioning is done. The service provisioning includes LSP establishment, LSP deletion and LSP modification.

Link management is used to manage TE links, e.g., maintain control channel connectivity, localize link failure, and so on.

Control information, e.g., signaling messages, routing messages, link management messages, is exchanged through the control channel. The control channel should be separated from the data channel as IETF recommended [10]. One of the good reasons for separation is that the control channel should not share the fate with the data channel. And it does not have to be the same physical medium as the data channel. For example, an OXC uses lambda to transport data, but uses an Ethernet link to transport control signals.

#### **2.4.1 Resource Discovery**

Local resource discovery is the procedure that a router takes to find out what resource it has for service provisioning.

When a node starts up, it goes through the neighbor and link discovery procedure, for example, by manual configuration or an automatic procedure. By combining the results, each node has a database about the local resource, for example, link capacity, wavelength, etc.

After the local resource discovery, each node uses the routing protocol to distribute its local resource. When a node receives other nodes' resources, it stores them in a database. Then, any changes to the resource will also be advertised by the routing protocol. Thus each node knows about the resource of the entire network.

#### **2.4.2 Enhancements in the Routing Protocol to Support GMPLS**

Conventional routing protocols are reused and enhanced with extensions to support GMPLS, e.g., OSPF with extensions [21], IS-IS with extensions [22]. They are used to discover network topology, distribute Traffic Engineering properties and GMPLS-specific features.

Here we introduce the extensions of conventional routing protocols to support unnumbered links, different interfaces, link protection type and Shared Risk Link Group distribution in GMPLS.

### Extensions to support unnumbered links

One of the fundamental issues in routing is addressing. Because of WDM, an optical fiber may have a number of channels. The IETF draft [14] suggests an addressing scheme: an IP address is used to identify a node (e.g., a router ID), and a “selector” is used to identify further fine-grain information within each node.

A numbered link means its interfaces are IP addressed. An unnumbered link means its interfaces are not IP addressed. In the optical network, optical fibers connect OXCs as point-to-point links. Point-to-point links need not to be numbered. In this case, the router (or an OXC) that connects an unnumbered link can assign a 32-bit identifier to the link. The identifier uniquely identifies the link within that router. So the identifier is locally significant. This local identifier is called the remote identifier from the point of view of the other OXC that is connected by the same unnumbered link. For example, OXC A and B are connected by unnumbered link L. OXC A assigns identifier L1 to L, which is a local identifier to A; OXC B assigns L2 to L, which is a local identifier to B. When the routing protocol exchanges the information between two routers, L1 is a remote identifier to B, and L2 is a remote identifier to A. The link can be uniquely identified globally by <router ID, (local) unnumbered link identifier> (see the example in Figure 1.6). Note that the router ID is always a 32-bit IP address.

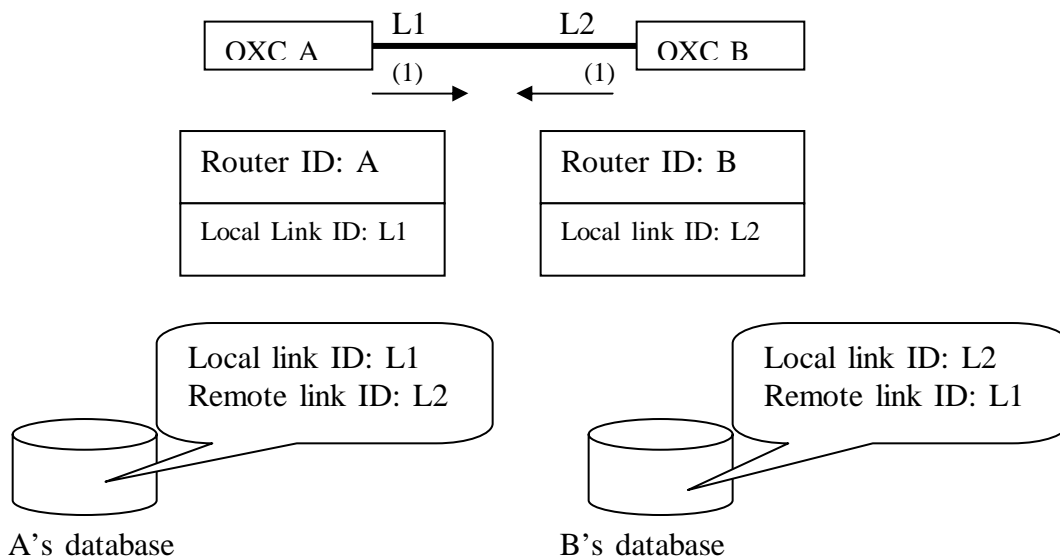


Figure 1.6: naming unnumbered link

It is assumed that an edge router that has physical connectivity to an OXC is able to provide optical-electrical data conversion. An edge router between the optical network and the IP network has interfaces that connect to OXCs and interfaces that connect to regular IP routers (see Figure 1.7). Assuming that the link F between the OXC and edge router is an optical fiber, and the link between the IP router and the edge router is a regular link (e.g., an Ethernet link). At the start-up, the edge router knows that the optical fiber F connects itself through interface I<sub>1</sub> to an OXC by neighbor discovery (e.g., by manual configuration, and see [23] for more about how a router discovers its

neighbors). And it knows that an Ethernet link connects itself through interface  $I_2$  to an IP router by neighbor discovery. When the edge router creates its routing adjacency relationship with its neighbors, it understands what parameters, options and protocol extensions it is going to use. Thus the routing protocol will send out advertisement messages carrying unnumbered link identifiers to identify link F, and it will send out advertisement messages carrying IP addresses to identify link L.

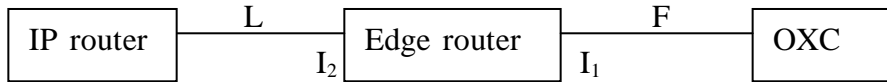


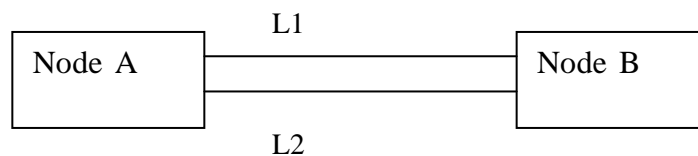
Figure 1.7: edge router knows about the links

### Extensions to support link protection type

If a link has a protection capability provided by the link layer, then such a link capability should be considered by the path computation component when calculating/selecting the path. The link protection type (e.g., 1+1 protection) is one of the traffic engineering properties of a link and it is distributed by the routing protocol. The link protection type does not have the same meaning when it is carried by signaling protocols as when it is carried by routing protocols, because it is from a different point of view. When the routing protocol distributes the link protection type for a given link, it means the link has the protection capability indicated by the link protection type. Let see what these link protection types are.

#### Extra Traffic

A link with type Extra Traffic means it is protecting another link or other links.



For instance, Link 1 and Link 2 connect Node A and Node B. Traffic is going through L2. If Link 1 is of type “Extra Traffic”, it is protecting L2, but there is no traffic going through L1 yet, or the traffic going through L1 is different from that going through L2.

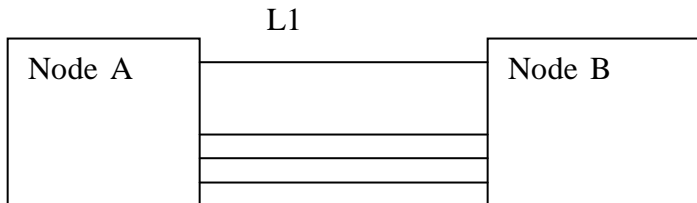
In Internet draft *Routing Extensions in Support of Generalized MPLS* [19], the sentence “The LSPs on a link of this type will be lost if any of the links it is protecting fail” means a link of this type will be activated when a link it is protecting fails. So any LSP that is on such a link will be preempted.

#### Unprotected

No link is protecting the link that is of type unprotected. If it fails, then the LSP is lost and so is the traffic.

### Shared

If the link is of type Shared, it means that there are one or more disjoint links of type Extra Traffic that are protecting this link.



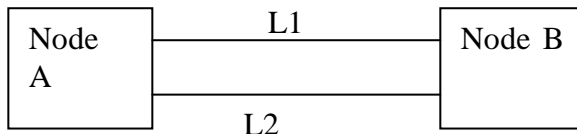
For instance, Link 1 is protecting one or more links, which is of type Extra Traffic. Other links that are protected by L1 are of type Shared – they share the protection relationship.

### Dedicated 1:1

If the link is of type Dedicated 1:1, it means that there is one dedicated disjoint link of type Extra Traffic that is protecting this link. For instance, in example of Type *Extra Traffic* (see above), Link 2 is typed *Dedicated 1:1*.

### Dedicated 1+1

If the link is of type *Dedicated 1+1*, it means that a dedicated disjoint link is protecting this link. However, the protecting link is not advertised in the link state database. So if the switchover occurs for a failure, the LSP is still there.



For instance, traffic is sent between two links: L1 and L2. The receiver takes the healthy one to accept user traffic. Link 2 and Link 1 both are of type *Dedicated 1+1*.

### Enhanced

A link of type Enhanced means it has a protection capability that is more reliable than *Dedicated 1+1*.

If the link information distributed by the routing protocol does not have the link protection type, it means it is unknown.

### Extensions to support Shared Risk Link Group

With the development of optical network, e.g., WDM, a number of links can have the same fate. Because they share the same physical resource, and if the resource is not available, then all these links are broken. For example, an optical fiber can contain a

number of links. Such a set of links constitutes a Shared Risk Link Group (SRLG) [19]. Based on different physical resource, a link may belong to multiple SRLGs.

For path protection/restoration, the links of the backup path must belong to different SRLG(s) from the ones of the working path. Therefore, the SRLG information is useful for the path computation component to compute the path.

### **Extensions to support different interfaces**

A link is connected to a node by an interface. GMPLS supports different types of interface, e.g., interface which is capable of packet switching, interface which is capable of lambda switching, etc. Different types of interface have different switching capabilities, and even same type of interface have different switching capabilities. The switching capability of the interface introduces a new constraint for path computation and signaling. In GMPLS LSP set up, a LSP must start and end at the same type of interface. So this information needs to be distributed onto the network.

The Interface Switching Capability Descriptor [24] describes the switching capability of an interface. The IETF draft *Routing Extensions in Support of Generalized MPLS* [19] defines the following interface switching types:

- Packet-Switch Capable-1 (PSC-1)
- Packet-Switch Capable-2 (PSC-2)
- Packet-Switch Capable-3 (PSC-3)
- Packet-Switch Capable-4 (PSC-4)
- Layer-2 Switch Capable (L2SC)
- Time-Division-Multiplex Capable (TDM)
- Lambda-Switch Capable (LSC)
- Fiber-Switch Capable (FSC)

If an interface is of type PSC, it means that the node receiving data over this interface can switch the received data on a packet-by-packet basis. An example is the Ethernet interface. Types PSC-1 through PSC-4 stand for different levels of capability. It means potentially an LSP starts and ends on PSC interface can also be nested into another LSP that also starts and ends on PSC interface assuming that the LSP interfaces have different switching capabilities. However the PSC types 1-4 has not been detailed in the draft yet.

If an interface is of type L2SC, it means that the node receiving data over this interface can switch the received frames based on the layer 2 address. An example is the ATM interface – based on ATM VCI/VPI to switch data.

If an interface is of type TDM, it means that the node receiving data over this interface can switch the received data based on the time slot. An example is the SONET interface.

If an interface is of type LSC, it means that the node receiving data over this interface can recognize and switch individual lambdas within the interface. An example is the interface of an OXC (or PXC) that can operate on an individual lambda.

If an interface is of type FSC, it means that the node receiving data over this interface can switch the entire contents to another interface. An example is the interface of an OXC (or PXC) that can operate on an individual fiber.

Besides the switching type, the Interface Switching Capability Descriptor also contains the maximum bandwidth for each priority (range from 0 to 7) that may be reserved on this link.

A link can be used to transport different data encoded in a different way, e.g., SONET, Lambda, Packet, etc. The data encoding method specifies this information in the Interface Switching Capability Descriptor.

So the Interface Switching Capability Descriptor contains three necessary pieces of information: (1) interface switching type, (2) max (reservable) bandwidth and (3) data encoding type. Optional information may be attached in the descriptor for some specific interface types, for example, if the interface is PSC, the Maximum Transport Unit should be specified. An example of an Interface Switching Capability Descriptor is like:

```
Interface Switching Capability = PSC-1
Encoding = Ethernet 802.3
Max Bandwidth[0] = 1.0 Gbps, for priority 0
```

When a node advertises its link information carrying the descriptor, the descriptor only describes the interface that connects the node originating the message. In the example in Figure 1.8, interface I and interface K connect the router A to other nodes. The Interface Switching Capability Descriptor (ISCD) originated by A only describes interface I and K, not the interface of another end of the link.

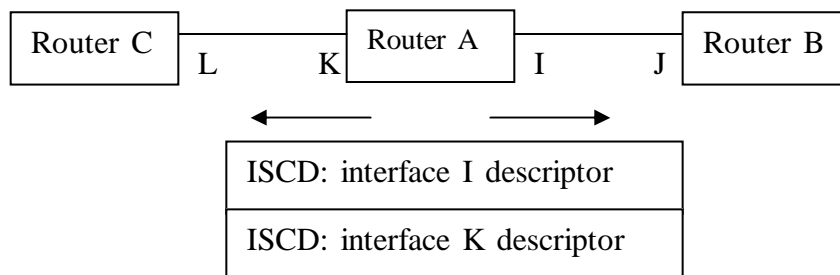


Figure 1.8: a router advertises the interface descriptor

### **Traffic Engineering properties**

Besides the above information, there are other TE properties that are distributed by routing protocols, e.g., maximum bandwidth, available bandwidth, etc. Because these TE properties are not specific for GMPLS, they will be introduced in the subsequent sections.

We are going to see how these extensions are implemented in OSPF as an example in the subsequent sections.

### **2.4.3 Enhancements in MPLS Signaling to Support GMPLS**

Signaling refers to exchange of information between involved components in the network required to provide and maintain service. GMPLS signaling provides LSP control (e.g., LSP set-up/release, LSP modification), and it may be used to reserve resources at the same time when LSP is being established. GMPLS signaling uses enhanced protocols CR-LDP [25] or RSVP-TE [26].

#### **Generalized Label Request and Generalized Label**

In the context of GMPLS, an LSP can be a mix of different types of link. For example, an LSP may have links that connects ATM switches, SONET switches, OXCs and others. And the label should take a different form. These forms of “label” are referred to as a *generalized label*.

In the GMPLS signaling, a node explicitly requests a label from its downstream peer when it needs one. The signaling message carries a label request, which should tell the downstream node enough information about the application environment of the desired label. The downstream node responds with a generalized label. It should contain enough information to allow nodes of the LSP to program their label forwarding tables.

Therefore, the signaling message should be extended to support the widening scope of GMPLS signaling. The label request message should include the following information:

- (1) LSP encoding type;
- (2) Switching type;
- (3) Generalized Payload ID (G-PID).

An LSP can be used to transport different data encoded in a different way, e.g., SONET, Lambda, Packet, etc. The LSP encoding types are defined in [27].

An interface connects a link to a node. The interfaces supported by GMPLS may have different switching capabilities, for example, packet-switch capable, lambda-switch capable, TDM capable, etc. These are named switching types in GMPLS signaling. A list of the switching types is defined in [24].

The Generalized Payload ID is an identifier of the payload carried by an LSP. Examples include lambda (using fiber), Ethernet (using fiber or lambda), etc. G-PID is defined in [27].



A generalized label has a variable length, which can model different types of “label”, e.g., wavelength, port, etc, in the context of GMPLS.

### **Bi-directional LSP setup**

There are a number of reasons [28] for using one signaling session to build a bi-directional LSP, instead of building two unidirectional LSP to do the same job. The advantages are obvious, e.g., the signaling overhead is less. From the restoration point of view, the delay to establish a bi-directional LSP to restore the service for a failed bi-directional LSP is less than the restoration delay for a unidirectional LSP. So the GMPLS signaling should be able to support bi-directional LSP set-up.

### **Label Set**

There are cases in GMPLS that result in label allocation trouble. For example, OXC A and OXC B are signaling neighbors for the set-up of a new LSP. OXC B (a downstream node) assigns label 10 to OXC A (an upstream node), which works as the outgoing label in A for forwarding data to B. But that label is not available in A (e.g., it does not have wavelength 10 at the interface to B). So the label set is defined in GMPLS signaling, which restricts the label range. For example, assuming that OXC A and OXC B both support GMPLS-RSVP-TE signaling, OXC A puts all the labels that are acceptable to A itself into the label set. The Path message carries the label set from A to B (from upstream to downstream). B can pick one of the labels in the set. However, if none of the labels in the label set is acceptable to B, B will generate an error and the path set-up will not continue.

### **Signaling Link Protection for LSP establishment**

During LSP signaling in GMPLS, label distribution protocols (RSVP-TE, or CR-LDP) may carry the link protection type. If the link protection type is carried, it means the LSP to be established requires link layer protection. The link protection type indicates what link protection capability is desired for the links constructing the LSP to be set up. The link protection type is one of the TE requirements (or a constraint) for an LSP, so the signaling for the LSP will not continue if the desired link protection cannot be provided. There are six link protection types defined by [27]. They have been specified in the previous section of this report. For example, the signaling protocol carries link protection type *Dedicated 1+1*, and it means the LSP to be established requires the link that has *Dedicated 1+1* protection.

### **Indication of the LSP role**

There are two LSP roles: primary or secondary (backup). The GMPLS signaling protocol carries a flag that indicates if the LSP being set up is primary or secondary. The resources allocated for a backup LSP are not used until the primary LSP fails. Because the resource allocation has priorities (carried by the signaling protocol), the resource allocated for a backup LSP may be used by an LSP that has lower priority until the primary LSP fails and the traffic is switched over to the backup. At that time, all the LSPs using the resource allocated for the backup LSP must be preempted.

#### 2.4.4 Path Computation

Traditional IP routing algorithms aim to find a path that optimizes a certain scalar metric (e.g. minimizes the number of hops), and such a method causes a number of network problems, e.g., network congestions, violation of network administration, etc.

Constraint-based routing algorithms set out to find a path that optimizes a certain scalar metric and at the same time does not violate a set of constraints. Such a path is called constraint-based path. It is the ability to find a path that does not violate a set of constraints that distinguishes constraint-based routing from conventional IP routing.

The constraints include QoS requirements, administrative policies, etc. Because we are studying the LSP protection/restoration mechanism, the constraint of interest is that the backup path must not share a link/node with the primary path except the initiator node and the terminator node. In particular, the information of Shared Risk Link Group and Link Protection Type are of interest to us. Note that the LSP role is for resource allocation and usage.

We need to compute a path to implement constraint-based routing. The path computation component in GMPLS control plane is used to do such a job. Path computation is used to select an appropriate route between two clients through the optical network for explicit routing.

In each node of the network, there is a database TE-LSDB that stores the information of all the links in the network, e.g., TE properties. This is the prerequisite for path computation. After all, we must know about the network before we calculate anything. Also, it means that the constraints we considered in the path computation are within the scope defined by the information in the TE-LSDB.

For a hop-by-hop routed LSP, there is no need to have path computation. When the signaling is done, it carries the desired Link Protection Type. Every node receiving the signaling message must honor the desired link protection for the LSP being established; otherwise, the signaling will not go through (see the subsequent section for more). Note that a hop-by-hop routed LSP cannot be the backup LSP, because there is no guarantee that the links/nodes traveled by such an LSP are not part of the primary LSP. The transit node is not supposed to keep track of the information about primary/backup LSP pairs, because there could be thousands of LSPs that go through a node.

Path computation is used to provide end-to-end LSP protection using the explicit-routed LSP (ER-LSP). If the primary LSP is an ER-LSP, then the backup LSP can be calculated following the primary LSP computation. If the primary LSP is a hop-by-hop routed LSP, and we know the nodes traveled by a hop-by-hop routed LSP, then we can also compute a path and use ER-LSP to create its backup. Otherwise, end-to-end LSP protection is not applicable.

Usually, constraint-based routing requires path computation at the LSP initiator node. This is because different LSP initiator node may have different constraints for a path to the same destination, and the constraints associated with a particular LSP initiator node are only known to that node. The reason is similar to source routing – the source determines the path.

The Shortest Path First (SPF) algorithm computes a path that is optimal with respect to some scalar metric. Many people (see [29]) propose that it is possible to modify the SPF algorithm in such a way that it can take into account the constraints. The algorithm is referred to as Constraint-based Shortest Path First (CSPF). There have been a number of proposals for CSPF, like [29]. The study of CSPF is out of the scope of this report, but a simple algorithm for CSPF is introduced to illustrate what CSPF is. It consists of three major steps:

- (1) Among all the links, exclude the ones that violate the constraints we defined.
- (2) According to the administration policy, map one (or more) link TE property as the scalar metric (cost) of the link.
- (3) Use the SPF algorithm to calculate the path.

Based on (1), we know that all the links we consider will not violate the constraints, and so will be the path. For example, the link color stands for an administrative constraint. If we want a path that is only within the “red” domain, then only the links with color “red” are considered. The user’s requirement is also a constraint – in fact, it is the most important one from the service point of view. If a user wants a path in which each link must have bandwidth 5Mb/s, then we do not consider all the links whose available bandwidth (the difference between the maximum bandwidth that may be reserved on this link and the bandwidth that has been allocated) is less than that.

With regard to path computation for LSP protection/restoration, the constraint is that the links traveled by the backup LSP must not belong to the same Shared Risk Link Group (SRLG) as the primary LSP. Therefore, after the computation for the primary LSP, all links belonging to the SRLG to which the links of the primary LSP belong are excluded (not considered).

In order to avoid the protection contention between LSP layer and link layer (see Section 5.1.2), [30] proposes that the Link Protection Type of the links traveled by the LSPs that construct the protection mechanism should be “unprotected”. Such a proposal is the second constraint that should be considered if we follow that proposal.

With regard to (2), we can take any of the TE properties or administrative distance.

Let us have an example. We will establish an LSP that requires T1 bandwidth (1.544 Mb/s), which travels from Node 1 to Node 5. In Figure 1.9, the link directly from Node 1 to Node 5 has only 1 Mb/s available; others have enough or more. So the link from Node 1 to Node 5 is excluded. Then we consider the available bandwidth as the metric.

The cost of a link is calculated by  $(10^8 / \text{available bandwidth})$ . The link from Node 1 to Node 2 has available bandwidth 10 Mb/s, so the cost is 10. In such a way, the metric of every link is calculated (see Figure 1.10). Then, using the SPF algorithm, the shortest path from Node 1 to Node 5 is (Node1, Node4, Node3, Node5).

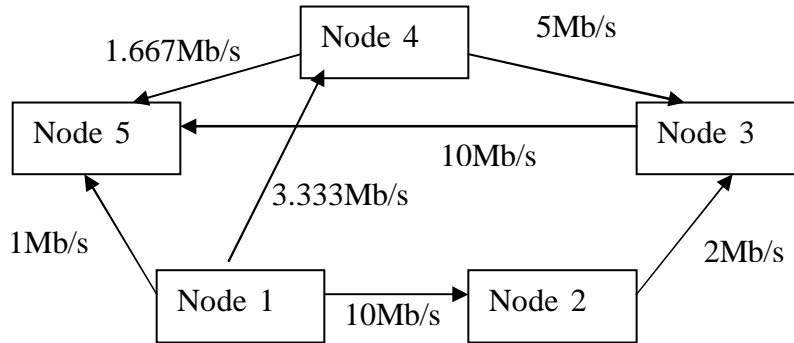


Figure 1.9: available bandwidth in the network

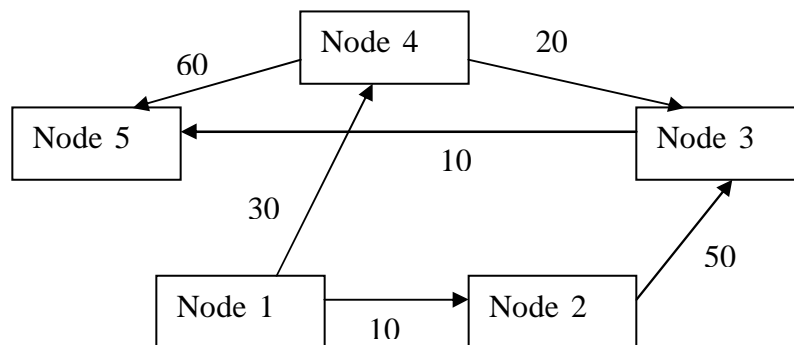


Figure 1.10: the metric of the links to be used by SPF algorithm

In general, path computation can be control-driven or data-driven. If the path computation is triggered by administrative control, e.g., the network administrator configures a path and requires the path computation for an ER-LSP, then the path computation is called control-driven. The data-driven path computation does not require administration. User data arrives at a node. In order to deliver the data, the node computes a path before signaling the LSP. Path computation is triggered by the data's arrival, and it is called data-driven. Using the control-driven mode, the path can be pre-calculated and even pre-established (before user data arrives), so it is faster in response to data delivery.

### 3. Overview of Path Protection/Restoration

With the development of networks, new technologies provide high bandwidth capacity. The ever-increasing bandwidth leads to a significant data loss if a failure cannot be recovered timely. Users and network service providers require network survivability. For example, real-time applications require very fast network recovery. No network service provider wants unprotected networks. On the other hand, transmission systems deployment gives chances to network failure, for example, telecommunication fiber cables share the same ducts of other utility transport media. Cable cuts are difficult to avoid.

Network survivability has been a hot research topic in the industry. Today, multiple layer protection/restoration is possible. The protection/restoration mechanism can be implemented in the link layer or in the IP/GMPLS layer. For example, the architecture of an IP-over-WDM node can be viewed logically as:

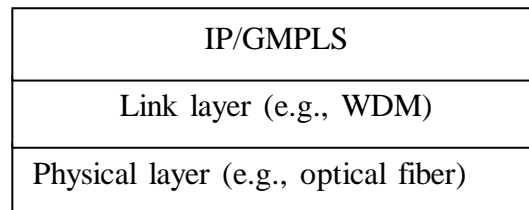


Figure 2.1: a logical view of the architecture of a GMPLS node

Protection/restoration mechanisms at the IP/GMPLS require relatively more time to recover, and using higher levels of recovery mechanisms may require more resources [31]. But there are limitations and disadvantages in the link layer protection, particularly in the optical network, e.g., complicated implementation, cost, instability due to duplication of functions, etc. It is still a challenge to implement recovery mechanisms at the WDM layer for the time being. Today a number of proposals have been studied in the industry to search for recovery mechanisms at the WDM layer, such as [32] and [33]. Furthermore, link layer protection cannot easily provide node protection [34]. The study of link layer recovery mechanisms is out of the scope of this report.

The motivation for using multiple layer protection is to provide the desired level of service in the most cost-effective manner [35]. With multiple layer protections, we need to prioritize them. The recovery mechanism that has higher priority is triggered first to recover failures. Usually, it is expected that lower layer recovery mechanism is closer to the failure, so it has higher priority. Also we need a coordination mechanism to avoid contention between different layer recovery schemes. One of the most popular coordination mechanisms is the hold-off timer. The hold-off time is the waiting time between the detection of a failure and taking MPLS-based recovery action. It allows time for lower layer protection to take effect [36]. If MPLS-based recovery is the only recovery mechanism desired, then the hold-off time may be zero. Assuming that we have SONET Automatic Protection Switch (APS) link protection, for example, within the hold-off time, GMPLS LSP path protection waits for the APS protection to switch. If the SONET APS succeeds protection within the hold-off time, then the hold-off timer is reset

and no further protection is needed. The original LSP can remain there. From this point of view, the link layer protection provides a means for LSP protection. Section 2.4.3 specifies how LSP signaling requires link layer protection when the LSP is being established. If the hold-off time expires, the LSP protection/restoration is triggered. The coordination mechanism introduces a tradeoff between rapid recovery and creation of a race condition where several layer protection mechanisms respond to the same fault.

GMPLS widens the application scope of MPLS, and people propose using GMPLS to build a unified control plane to manage all kinds of network nodes [14]. The GMPLS LSP protection/restoration has been an important recovery mechanism for network survivability.

Differently from traditional IP networks, MPLS networks establish label switched paths (LSPs) before data forwarding occurs. This potentially allows MPLS networks to pre-establish protection (backup) LSPs for working LSPs, and achieve better survivability than traditional IP networks.

Here we introduce what we need for the LSP protection/restoration mechanism in GMPLS networks.

- (1) A method for computing the working and protection paths;
- (2) A method for working and protection path signaling;
- (3) A fault detection mechanism;
- (4) A fault localization and notification mechanism to localize the fault and convey the information;
- (5) A recovery mechanism to move the traffic over from the working path to the protection path or to reroute the fault;
- (6) A repair detection mechanism to detect the original working path is fixed;
- (7) An optional switchback or restoration mechanism to restore the traffic to the original working path.

Item (7) is optional and it is not time-sensitive. In some cases, it may not be desirable. For example, switching the traffic back to the original working path can disrupt the traffic (even for a very short time). It may not be desired under the user requirements. Item (6) may not be necessary in some cases. For example, if (7) is not wanted, then (6) is not needed.

Item (1) is implemented by the path computation component. For example, it uses CSPF to compute a path and selects the working and protection path. Usually it is proprietary. The path computation considers the Traffic Engineering properties of the network, administrative constraints and user requirements to calculate the backup and working path. For example, if both Link L and K share the same physical resource (e.g., they exist in the same optical cable), then either L or K should be considered in a particular working path computation and its backup.

As we introduced in the last sub-section, the GMPLS signaling protocols carries the link protection information, which can allow the nodes on the network to identify the working and backup path.

Traditional methods to monitor the health of data links may not be useful any more. For example, pure optical switches may not allow these methods to check the bit-rate, format or wavelength. Fault detection should work at the layer closest to the failure in order to achieve quick response. In optical network, this should be located in the physical layer (e.g., optical layer). For example, one method of fault detection at the optical layer is detecting the loss of light (LOL). Using software can also detect a faulty link/node, and it will be introduced in the subsequent section. However, fault detection at the physical layer provides fast and reliable solution, and it is preferred if it is applicable.

The optical network has its own character in failure. When one link is broken, e.g., a fiber cut, all the downstream nodes (in terms of data flow) can detect loss of light. Therefore, we also need a method to localize the failure. The Link Management Protocol provides a method, which will be introduced in the subsequent section.

Both GMPLS signaling protocols [26] and [25] are being extended to provide methods to support LSP protection/restoration. For simplicity, we use the term RSVP-TE to refer to [26] and CR-LDP to [25] from now on.

There are a number of objectives for the LSP protection/restoration mechanism. The LSP protection/restoration mechanism should

- (1) optimize the use of resources;
- (2) provide fast recovery and minimize the disruption to data traffic of any failure;
- (3) minimize degrading the traffic and preserve the constraints on the traffic after switchover;
- (4) minimize the recovery overhead (be simple);
- (5) be cost-efficient.

At the end of our discussion, we will see that some of the above objectives are conflicting. There is a trade-off between them. It is impossible to achieve all of these objectives at the same time, and the choice depends on what the user wants and what is the network administration goal.

## **4. Multiple Protocols Contribute to GMPLS LSP Protection/Restoration**

### **4.1 OSPF Extensions**

The current routing protocols OSPF and IS-IS are extended to support Traffic Engineering and GMPLS. Here we take the popular OSPF as an example to see how it works.

The OSPF protocol is re-used to distribute information to support Traffic Engineering and GMPLS features. Two types of extensions have been added to the OSPF: TE extensions and extensions for GMPLS. The former is named OSPF-TE, which distributes TE properties over the network. The latter is referred to as GMPLS-OSPF, which distributes extensions dedicated to support GMPLS.

In the OSPF protocol, the message describing the local link information that is flooded throughout the network is named Link State Advertisement (LSA). A new LSA - TE LSA is defined to support Traffic Engineering and GMPLS (see [37] for more information).

The Type-Length-Value (TLV) structure (see Figure 3.1) is used as the payload in the TE LSA. The Type specifies the type of the data; the length specifies the length of the whole TLV structure, and the Value describes the information regarding to Traffic Engineering and GMPLS support.

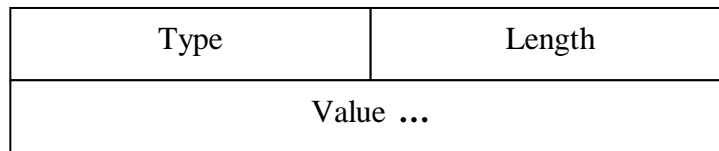


Figure 3.1: the TLV structure

The TLV structure can be nested, for example, sub-TLVs are carried as the value in the higher-level TLV. So it is extendable, which is good for future development. There are two TLVs: router address TLV and link TLV.

### **Router address TLV**

The router address is the router ID of the node that advertises the LSA. The TE LSA must carry a router address TLV. It is type 1, the length is 4, and the value is the 4-octet IP address.

### **Link TLV**

The link TLV contains information about the link. And it consists of a set of sub-TLVs, each of which describes a piece of particular information about Traffic Engineering or GMPLS features. The information of these sub-TLVs are introduced in the subsequent sections. The Link TLV is type 2 and the length varies.

OSPF does not process the contents of the TE LSA.

#### **4.1.1 Introduction to Traffic Engineering Extensions to OSPF (OSPF-TE)**

When a router starts, it discovers the information about its own links (interfaces) – the links connecting the router to networks (or other routers). Then the routing protocol is used to advertise the information to other routers. The information is passed around from router to router. Ultimately, every router has identical information about the network and



the information is stored in a database named Link State Database (LSDB). Each router will independently calculate the best path to other nodes in the network using a path computation algorithm. For example, the popular OSPF protocol uses Dijkstra's Shortest Path First (SPF) algorithm to come up with a SPF tree, which serves as a map for data routing (see Figure 3.2). Then according to routing policies, an appropriate route is selected and put into the routing table.

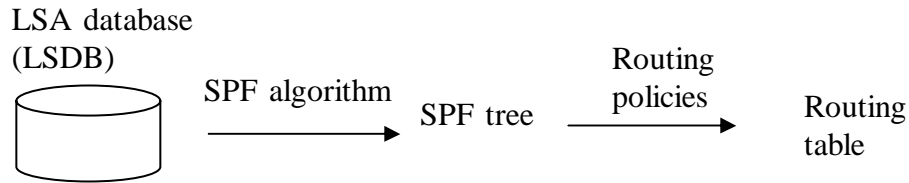


Figure 3.2: from LSDB to an appropriate route

Conventionally, the information of the links includes the status of the links (e.g., up/down), metric (cost), etc. The information does not support Traffic Engineering. For example, the metric is assigned to routes as a means of ranking them from the most preferred to the least preferred. The calculation of the metric is static. The bandwidth metric used in Cisco routers is calculated as:  $\text{metric} = 10^8 / (\text{link bandwidth})$ . Thus a higher-bandwidth path is always preferred over a lower-bandwidth path. But what if a T1 link of the preferred path is heavily loaded with traffic and a 64k link is lightly loaded?

Because the TE properties (e.g., bandwidth availability, administrative constraints) are not provided or considered in conventional routing protocols, the routing decision does not support Traffic Engineering.

Relying on the current routing protocols, TE properties are added into the messages that are flooded throughout the network. In IETF, the draft OSPF-TE [37] proposes the following TE properties that should be considered to support Traffic Engineering, and they rely on the OSPF opaque LSA advertising mechanism to distribute the TE properties. Each of the following 9 items constructs a sub-TLV in the link TLV of the TE LSA. Note that they are optional except the first two sub-TLVs: Link type and Link ID.

- 1 - Link type
- 2 - Link ID
- 3 - Local interface IP address
- 4 - Remote interface IP address
- 5 - Traffic engineering metric
- 6 - Maximum bandwidth
- 7 - Maximum reservable bandwidth
- 8 - Unreserved bandwidth
- 9 - Resource class/color

**Link type**

It specifies if the link is (1) point-to-point or (2) multi-access link. For the time being, only point-to-point link is completely supported.

**Link ID**

The Link ID identifies the remote end of the link. For point-to-point links, this is the Router ID of the neighbor.

**Remote Interface IP Address**

It specifies the IP address of the neighbor's interface corresponding to this link. For unnumbered links, this is the link remote identifier (see Section 2).

**Local Interface IP address**

It specifies the IP address(es) of the interface corresponding to this link. If there are multiple local addresses on the link, they are all listed in the appropriate structure of a routing message. For unnumbered links, this is the link local identifier (see Section 2).

The local and remote interface IP addresses identify the parallel links between two nodes.

**Traffic Engineering Metric**

A metric is a variable assigned to routes as a means of ranking them from best to worst or from most preferred to least preferred. The Traffic Engineering metric specifies the link metric for traffic engineering purposes. This metric may be different than the standard OSPF link metric.

**Maximum Bandwidth**

It specifies the maximum bandwidth that can be used on this link from the LSA-originating router to its neighbor. For example, a T1 link has maximum bandwidth 1.544 Mb/s, an OC-48 link has around 2.5 Gb/s.

**Maximum Reservable Bandwidth**

It specifies the maximum bandwidth that may be reserved on this link in the direction from the LSA-originating router to its neighbor. Note that this may be greater than the maximum bandwidth (the link may be oversubscribed). For example, an OC-48 link may be configured to have maximum reservable bandwidth 2.75 Gb/s (10% oversubscribed).

**Unreserved Bandwidth**

It is the difference between the Maximum Reservable Bandwidth and the bandwidth that has been reserved. There are eight priority levels (from 0 to 7) of unreserved bandwidth. This information specifies the unreserved bandwidth of each priority level. Priority 0 is the highest.

**Resource Class/Color**

It specifies administrative group membership for this link, in terms of a bit mask. A link may belong to multiple groups - if so it has multiple bit masks.

A node advertises the TE-LSA whenever one of its own links gets the TE properties updated. The routers that receive these TE-LSAs store them in a database that is named TE Link State Database (TE-LSDB). The TE LSDB is synchronized across all nodes supporting OSPF-TE within an area. So each node in that area has an identical view of the TE properties of the network. The path computation component of the control plane can use the information provided by TE LSDB to compute a path that meets a user's requirements and the traffic engineering goals (see Figure 3.3).

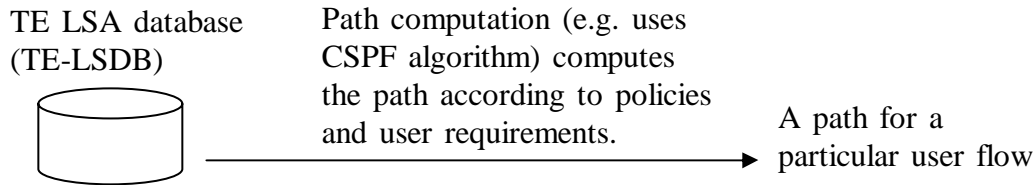


Figure 3.3: from OSPF-TE to a path

Like the regular links, FA-LSPs (an LSP is advertised as a link in the network - see the Section *Traffic Engineering* before) also have the TE properties we just introduced. They are also stored in the database TE-LSDB. This information is also used by the path component to compute a path. As examples, here we list some of the TE properties of a FA-LSP (see [20] for more).

- (1) Link type: an FA-LSP must be a “point-to-point” link;
- (2) Local and Remote interface address: if the FA-LSP is to be numbered, then the local interface IP address is the head-end address of the FA-LSP link. And the remote interface IP address is the address of the ending node of the FA-LSP;
- (3) Maximum Bandwidth (also named Maximum LSP Bandwidth): It specifies the maximum bandwidth that may be reserved on this LSP. Therefore, it is like the Maximum Reservable Bandwidth of a link.
- (4) Interface Switching Capability: it is the Interface Switching Capability of the first link of the FA-LSP.

As it is introduced, the above TE properties are carried by the TLV structure within the TE LSA and distributed by OSPF.

#### 4.1.2 Extensions to OSPF for supporting GMPLS

The following information is needed to support GMPLS: (1) unnumbered link identifier; (2) Link Protection Information; (3) Shared Risk Link Group (SRLG) Information; (4) Interface Switching Capability Descriptor. They also rely on the TE LSA of OSPF to be distributed into the network.

##### 4.1.2.1 Unnumbered link support in OSPF

How unnumbered link is supported has been introduced in Section 2. In OSPF, the 32-bit unnumbered link identifier (local and remote) is simply put into the value field of the TLV

structure. The type is 11. If the remote identifier is unknown (e.g., at the router start-up), then it is 0. Carried by the TE LSA, the unnumbered link identifier is advertised.

#### 4.1.2.2 Shared Risk Link Group (SRLG)

The SRLG is also a link property and it is advertised by the link sub-TLV. The SRLG is specified by a 32-bit word, contained in the Value field of the sub-TLV structure. The sub-TLV type is 16. If a link can belong to multiple SRLG, then all of them are listed in the sub-TLV structure and the order is irrelevant. An example is shown in Figure 3.4.

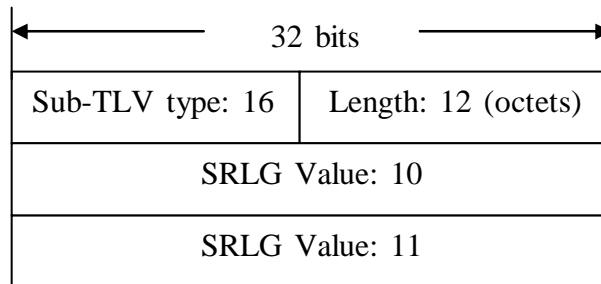


Figure 3.4: SRLG sub-TLV

#### 4.1.2.3 Link Protection Type

The link protection type can be considered by the path computation component to compute a path and it is distributed throughout the network. There are six link protection types (See Section 2.4.2 of this report for what they are.):

- (1) Extra Traffic;
- (2) Unprotected;
- (3) Shared;
- (4) Dedicated 1:1;
- (5) Dedicated 1+1;
- (6) Enhanced.

If the routing protocol does not distribute the link protection type for a link, then the protection attribute of that link is unknown.

The link protection type is encoded in a sub-TLV of the link TLV. The sub-TLV type is 14 and it has 4 octets (see Figure 3.5). But only the first octet is used. The first octet is used for indicating protection types and the other octets are reserved. The first octet may contain the following value to indicate the link protection type:

- 0x01 Extra Traffic
- 0x02 Unprotected
- 0x04 Shared
- 0x08 Dedicated 1:1
- 0x10 Dedicated 1+1
- 0x20 Enhanced
- 0x40 Reserved
- 0x80 Reserved

Protection type	reserved
-----------------	----------

Figure 3.5: the Value field of the sub-TLV for link protection type

#### 4.1.2.4 Interface Switching Capability Descriptor

The interface switching capability is encoded by a sub-TLV (type 15) of a link TLV. The field contains one of the following codes. And each code signals the correspondent type.

##### Code Type

1	Packet-Switch Capable-1 (PSC-1)
2	Packet-Switch Capable-2 (PSC-2)
3	Packet-Switch Capable-3 (PSC-3)
4	Packet-Switch Capable-4 (PSC-4)
51	Layer-2 Switch Capable (L2SC)
100	Time-Division-Multiplex Capable (TDM)
150	Lambda-Switch Capable (LSC)
200	Fiber-Switch Capable (FSC)

The code is not consecutive, as it allows for future extension.

#### 4.2 Link Management Protocol (LMP)

Neighboring nodes may run the Link Management Protocol (LMP) [38] for link management. With the development of optical networks, nodes include photonic switches (PXC's), optical crossconnects (OXC's), routers, switches, add-drop multiplexors, WDM systems and so on. LMP support any type of nodes. And LMP supports TE links.

The link multiplexing capability has an effect on how to do the link management, e.g., resource allocation. To allow interworking between links with different multiplexing capability, sub-channels of a component link should be able to be configured as a data link. For example, several Ethernet links are multiplexed into an OC-12 link, which is connected to a node. The node should allow each Ethernet link to be configured as a data link. So that link management on each Ethernet link is possible if required.

To run LMP, a control channel must be established between the node pair. The control channel should be separated from the data channel [10]. And, the node pair can communicate bi-directionally at least through one of the control channels. If so, then an LMP adjacency can be formed between the two nodes. Multiple active control channels are possible in an LMP adjacency, and the control channel ID (CCID) is used to identify each one.

LMP messages are encoded as data in IP packets, and it runs directly over IP except for the LMP Test message. The LMP Test message is sent over the data links (in-band) for

link connectivity verification. So optionally it is limited by the transport media, e.g., not necessarily encoded as data in IP packets.

LMP functions are: control channel management, link property correlation, link connectivity verification, and fault management.

(1) **Control channel management** is used to establish and maintain control channels between LMP adjacent nodes. The control channel can be used to exchange routing, signaling, and other control messages.

To establish the control channel, the IP address for the far-end of the control channel must be known (e.g., by configuration). A node sends a LMP Config message to its neighbor, which contains parameters, e.g., the LMP keep-alive interval. The receiver of the Config message must reply an acknowledgement. If both sides agree on the parameters, the control channel is established. After that, the LMP keep-alive message is sent periodically to maintain the control channel.

After two neighboring nodes successfully establish the control channel, control messages can be exchanged through the control channel. Examples of these control messages may be label distribution information implemented by RSVP-TE, network topology and state distribution information implemented by OSPF-TE, fault management implemented by LMP, and so on.

(2) **Link property correlation** is used to synchronize the properties of the TE link and verify the configuration. An example of TE link is shown in following figure. LSP is taken as a TE link by Node1 and Node3, which is constructed by link (A, B) and link (C, D). Link (A, B) or link (C, D) is called a data link.

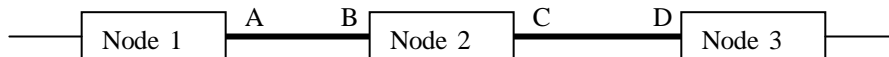


Figure 3.6: An LSP as a TE link starting from Node1 to Node3

After the LSP is established by the signaling protocol, LMP may be used to synchronize the properties of the TE link. So, Node1 may send a LMP LinkSummary message to Node3, which is constructed by LMP objects as:

```
<LinkSummary Message> ::= <LMP message header><Message ID>  
                           <TE Link><Data Link (A, B)><Data Link (C, D)>
```

Within each Data Link object, sub-objects may contain information about link reservable bandwidth, wavelength if there is any, interface switching capability such as interface A for data link (A, B).

The receiver of LinkSummary message must verify that the information obtained from the message makes sense and matches the information that is stored in the routing database or configuration inventory. For instance, the interfaces A and D must be of the same interface switching capability type in the example shown in Figure 3.6. The receiver of a LinkSummary message must reply an acknowledgement, which reports the correctness of the TE link properties.

(3) **Link connectivity verification** is used to verify the physical connectivity of the data links between the nodes.

In the example shown in Figure 3.6, Node1 and Node3 may exchange LMP Test messages between interface A and D through link (A, B) and (C, D) to verify the physical connectivity of the TE link on a periodic basis. The verification messaging must be transported by the data-bearing channel, not the control channel.

(4) **Fault management** provides a fault localization procedure. Because the LMP fault management is within the scope of this report, let us discuss it in detail.

The Link Management Protocol introduces a fault localization procedure to localize failures. It can localize the path failure by quickly reporting the status of one or more data link. It is designed to work for both unidirectional and bi-directional LSPs.

During the Link Property Correlation, both LMP-capable nodes can signal whether they support LMP fault management. If they do, then LMP fault management messaging becomes one of the control signals between these two nodes.

In optical networks, e.g., nodes are PXC's in the network, if one of the data links fails, then all the downstream nodes (in terms of data flow) may detect the failure due to the nature of light, e.g., loss of light. The LMP fault management requires each node that has detected the failure to send a LMP ChannelStatus message to the upstream node. This ChannelStatus message can report all the broken channel/links together. The upstream node must acknowledge the message by a LMP ChannelStatusAck message. Then the upstream node checks if there is any local data link failure, for example, it checks if the input side has any signal. If the input side is working fine, the failure is localized; otherwise, the node will continue sending LMP ChannelStatus messaging upstream. After the local checking, the upstream node must send a ChannelStatus message to the downstream node to report the status.

On the other hand, after the downstream node receives the ChannelStatusAck, it expects a ChannelStatus from the upstream node. If it receives no ChannelStatus, it should send a ChannelStatusRequest to solicit the message.

The time-sequence diagram in Figure 3.7 outlines how it works. Let us suppose that Node 2 is the downstream node relatively to Node 1 (in terms of data flow). Node 2 detects a failure.

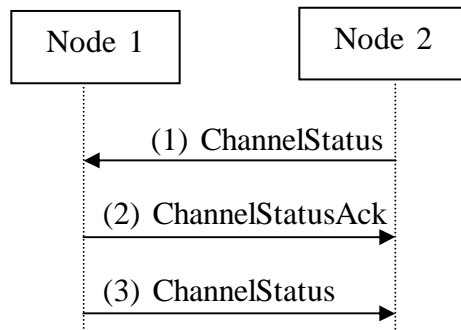


Figure 3.7: channelStatus messaging

When the fault is localized, the upstream node which connects the failed link should trigger the signaling to get protection/restoration. And it does not perform LMP ChannelStatus messaging to upstream nodes any more.

Let us have an example to see how it works in a pure optical network. There are three PXC's in the example shown in Figure 3.8. An LSP travels the data links of these three nodes. The control channel is out-of-band. Assuming that the data link through which the LSP with the flow direction from PXC 1 to PXC3 is failed. Both PXC 2 and 3 can detect the failure. For simplicity, only one direction of the LSP is shown.

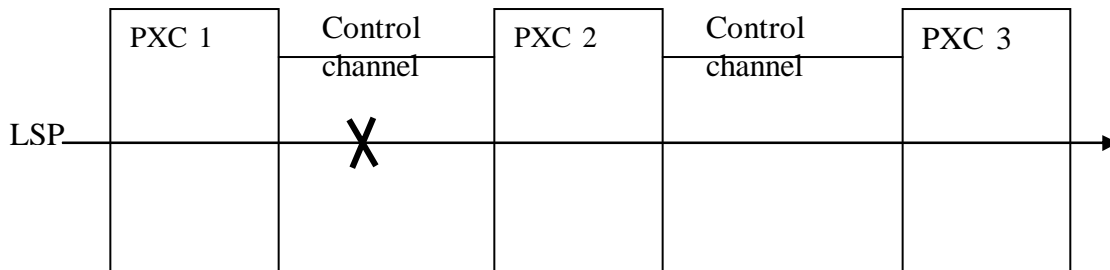


Figure 3.8: LMP fault management localizes the fault

PXC 3 sends the LMP ChannelStatus message to PXC 2, which acknowledges with a ChannelStatusAck. PXC 2 locally finds out that there is no input signal and the failure is propagated from upstream. So it tells PXC 3 also by a ChannelStatus message. Meanwhile, PXC 2 sends another ChannelStatus message to PXC 1, which tells PXC 1 that no signal comes in. PXC 1 replies with a ChannelStatusAck. PXC 1 locally finds out that the input is fine. So it sends PXC 2 a ChannelStatus message, which tells PXC 2 that it is clear. Thus PXC 1 has localized the failure. After that, the recovery will be triggered, for example, signaling starts to establish a reroute. Section 5 will specify the recovery mechanisms in detail.

If the failure affects both directions of the LSP, e.g., a fiber cut, then the same procedure is performed on each direction.



### 4.3 GMPLS Signaling

There are two major label distribution protocols to perform GMPLS signaling: RSVP-TE with extensions and LDP with extensions.

#### 4.3.1 GMPLS signaling: RSVP-TE with extensions

Traditional RSVP (RFC2205) provides a means for an application to communicate its QoS requirements to an Integrated Services Internet infrastructure. RSVP is a control protocol that signals QoS requirements on behalf of a data flow. Before data delivery occurs, RSVP establishes a resource reservation for a simplex (one way) flow along its path. A simplex flow is a unidirectional flow traveling from its source to its destination. To allow duplex (two-way) communication, we need RSVP to reserve resource twice – one for each direction. RSVP consults a routing table in a router for the next hop. RSVP relies on IP or UDP for message transport.

RSVP must carry the following information:

- Information for flow identification, so that the flows with particular QoS requirements can be recognized within the network. This may include sender IP address, receiver IP address, port numbers and so on.
- Traffic specification and QoS requirements.

RSVP carries the information from the source host to the destination host along the router/switch on the path. There are two basic messages in RSVP: PATH and RESV messages. A PATH message travels from the sender to the receiver and include traffic specification and classification information provided by the sender. The PATH message identifies the path from the sender to the receiver and it collects status about the resource along the path. When the PATH arrives at the receiver, the receiver sends back a RESV message back toward the sender along the reverse of the path. The RESV message communicates with every router to make a resource reservation. See the following figure for PATH/RESV messaging.

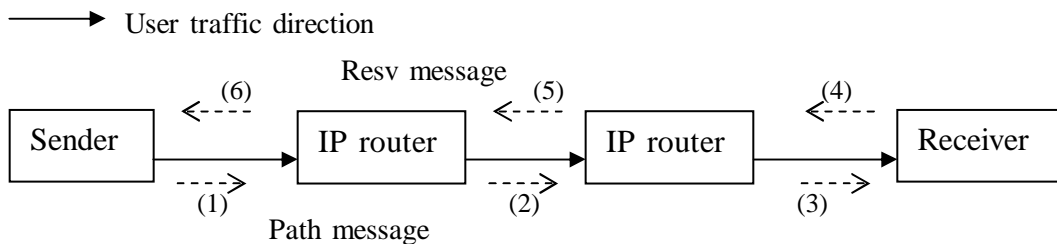


Figure 3.9: RSVP signaling to reserve resource

Each router along the path creates a software record (software state) for the particular flow, which keeps the flow classifier, QoS requirements, next hops, previous hops and other related information. These records have a timer, which means these software states will be removed after some time-out. So after some time period, the PATH message is

transmitted and the RESV travels the reverse path – the process repeats on a regular time interval basis. This is called refresh messaging, which keeps the software states and the router can continue providing QoS to the flow.

Extensions have been added to RSVP (RSVP-TE) to support label distribution for LSP signaling in MPLS. To establish an LSP, the sender node, with respect to the path, creates an RSVP Path message which contains a LABEL\_REQUEST object. The LABEL\_REQUEST object indicates that a label binding for this path is requested. A SESSION\_ATTRIBUTE object is introduced to provide additional control information such as setup and hold priorities, local protection and so on. The RSVP-TE Path message carries this object during LSP signaling. When the Path message arrives at the destination node of a LSP, the node responds to the LABEL\_REQUEST object with a LABEL object in its RSVP-TE Resv message. If the node is not the sender node, it allocates a free label and puts it into the LABEL object. And the Resv message is sent to the upstream node. The node that receives a Resv message with a LABEL object will use this label as the outgoing label in the forwarding entry of its forwarding table. It also allocates a free label for the upstream node, and puts it into the LABEL object attached to the Resv message. The Resv message is sent upstream again. Such a label distribution procedure repeats until the Resv message arrives at the sender node. The LSP establishment is done. The sender node has some criteria to classify different traffics and puts the predefined traffic into the appropriate LSP. In the example shown in Figure 3.10, the sender node will attach label 3 to all the packets before it forwards the packets out. When the packet arrives at the transit IP router, the label is replaced by 6 and then forwarded again. Such a label swapping procedure repeats on each node and the packet finally reaches the destination.

For label distribution, the LABEL\_REQUEST and LABEL objects are mandatory, but other objects defined in RSVP-TE are optional, e.g., the SESSION\_ATTRIBUTE mentioned above.

Note that because the label distribution is done with RSVP, each router can associate the resources with the LSP during LSP signaling. Therefore, resource reservation can be done in the meanwhile.

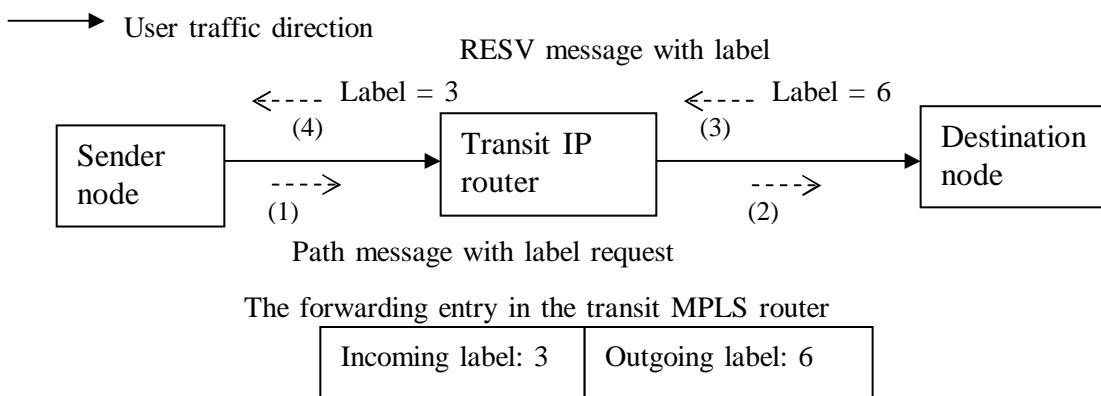


Figure 3.10: RSVP-TE signaling to distribute labels for establishing LSP

In each intermediate node, RSVP-TE consults the local routing table for the next hop. The LSP established in this way by RSVP-TE is named hop-by-hop routed LSP. Signaling in this way does not meet the requirements of many applications, for example, traffic engineering. So, the Explicit Route Object (ERO) is added to RSVP-TE to support the explicitly routed LSP (ER-LSP), which is similar to source routing. This object allows the path taken by RSVP-TE messaging to be pre-determined by the source. With ERO, the ingress node of the LSP can define which transit node the LSP will travel to reach the destination (egress node of the LSP). And the ER-LSP can be routed away from network failures, bottlenecks, or congestion.

The ERO is carried by the RSVP Path message. It contains a sequence of IP prefixes or a sequence of Autonomous Systems. The ERO tells the routing mechanism where to forward the Path message. We consider the following an example shown in Figure 3.11.

R1 is going to establish an explicitly routed LSP (R2, R3, R4, R5). R1 constructs the object ERO to have the sequence of nodes - R2, R3, R4 and R5. And each node can be represented by an IP address prefix. Then R1 creates the RSVP PATH message carrying the ERO as well as the LABEL\_REQUEST object. Before the message is sent out, R1 checks the top of the ERO, and ERO tells R1 the next hop is R2. R1 sends it to R2. R2 looks at the top of the ERO and finds itself is on the top. R2 looks at the next one, which is the IP address prefix for R3, and takes it as the next hop for the message. R2 removes the top IP address prefix that is one of its interfaces through which the message comes in, before it forwards the message. R3, R4 and R5 follow the same algorithm as R2 does. When R5 receives the PATH message, ERO only has one prefix, which is one of the interfaces of R5. Note that a RSVP state has been created on every router along the path.

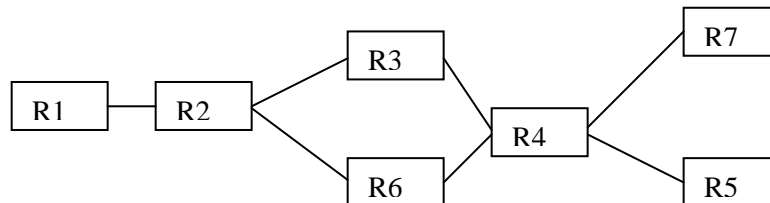


Figure 3.11: ER-LSP from R1 to R5

In order to respond to the LABEL\_REQUEST object, the R5 constructs a RSVP RESV message along with the LABEL object. The message can be forwarded to R4 by R5. R4 updates the LABEL object and further forwards the message to R3. The message follows the RSVP state that the PATH message has created along the routers R4, R3, R2 and finally reaches R1. Thus a LSP is created. Note that an intermediate router may not be able to tell the difference between a label for an established, explicitly routed LSP and

one for a hop-by-hop routed LSP, as it does not need to make this distinction during programming the data forwarding plane.

If the ERO specifies every node of the LSP or every autonomous system traveled by the LSP, then the LSP is called “strictly” explicitly routed. If the ERO specifies some nodes or some autonomous systems traveled by the LSP, then the LSP is called “loosely” explicitly routed.

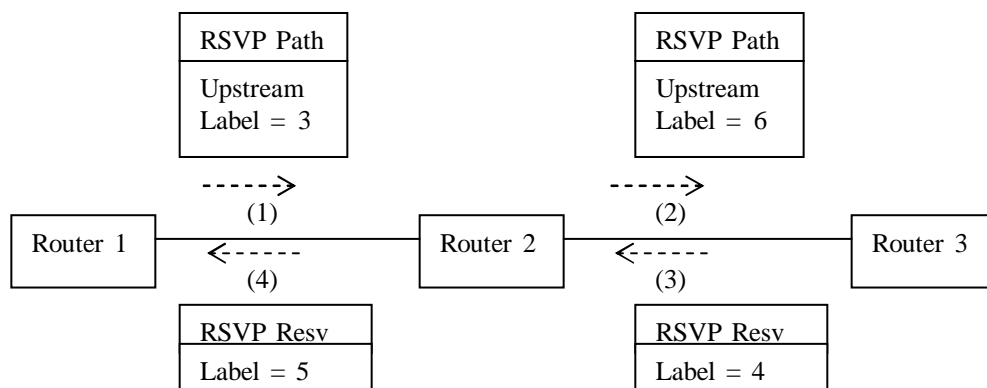
GMPLS extends MPLS to support multiple different interfaces. RSVP-TE is also extended to support GMPLS signaling. The label request object and label object must be generalized (see Section 2.4.2 of this report). In [39], the Generalized Label Request Object (carried by the Path message) and Generalized Label Object (carried by the Resv message) are defined. The Generalized Label Request Object allows different transit nodes with different data links to allocate labels.

When the Path message carrying the Generalized Label Request Object arrives at a node, the node makes sure the label request information (including the switching type, LSP encoding type and generalized payload ID) must be satisfied by the interface through which the traffic comes (incoming interface), the node itself and the interface through which traffic gets forwarded (outgoing interface). The node itself and the interfaces through which the traffic is transmitted should be able to support the LSP encoding type. The incoming interface should be able to support the switching type. Note that the label switched path (LSP) can be established only between (or through) interfaces of the same switching type. Usually only the egress will check the generalized payload ID (because the payload is transparent to transit nodes). If the egress does not support the payload, the LSP cannot be established. In all of these cases, a RSVP-TE PathErr message is generated.

There is no internal structure within a label. If we want nested LSPs (an LSP within another LSP), each LSP must be established separately.

The RSVP-TE Resv message carries the generalized label upstream along the reverse path set up by the Path message. The node that receives the Resv message must verify that the label is acceptable. In some situations, the label assigned by the downstream node could not be available, for example, an optical cross-connect does not have the wavelength to model the label. If the label is not acceptable, the node will generate a RSVP-TE ResvErr message.

In GMPLS-RSVP-TE, a procedure for bi-directional LSP set-up is introduced. The procedure is added to the establishment of a unidirectional LSP. The `Upstream_Label` object is defined in [39] and it is carried by the RSVP-TE Path message. This object is similar to the Generalized Label object. It contains a generalized label that is allocated by the upstream node and used by the downstream node for label swapping. An example is shown in Figure 3.12. The node that receives the `Upstream_Label` must verify the label is acceptable.



The forwarding entry in Router 2 for direction from R1 to R3

Incoming label 5	Outgoing label 4
------------------	------------------

The forwarding entry in Router 2 for direction from R3 to R1

Incoming label 6	Outgoing label 3
------------------	------------------

Figure 3.12: bi-directional LSP set-up using RSVP-TE

To support explicitly routed LSP in the context of GMPLS, just the IP address or the identifier of an autonomous system may not be adequate. For example, the LSP set-up needs to concatenate two LSPs to form an LSP at the edge of two different networks (e.g., an optical network and an IP network). There may be a number of wavelengths in a fiber (a link), and a particular wavelength (a label) is needed. The ingress of the LSP needs to specify the particular label (wavelength). So to support GMPLS signaling, a Label subobject is defined, which follows the IP address or the identifier of an autonomous system in the ERO. The Label subobject allows the ingress of the LSP to specify a particular label of a data link.

To improve network survivability, the protection information is considered in GMPLS signaling. It includes

- (1) link protection type;
- (2) indication of whether the path is primary or backup.

The link protection type indicates what link protection capability is desired for the links constructing the LSP to be set up (see Section 4.1.3.3 for the link protection types). During LSP signaling in GMPLS, label distribution protocols (RSVP-TE, or LDP) may carry the protection information. The link protection type in the protection information is one of the TE requirements (or a constraint) for a LSP to be set up. So the LSP set-up will not continue if the desired link protection cannot be provided.

### Signaling a hierarchical LSP

GMPLS supports interfaces that have different switching capabilities. The Interface Switching Capability Descriptor describing the capability is distributed by the routing protocol throughout the network (see the section *Enhancements in the Routing Protocol to Support GMPLS*), and each node stores this information in the TE link state database (TE-LSDB).

An edge node is the one that connects two different networks constructed by different nodes, for example, an optical switch that has interfaces providing SONET signals and interfaces providing WDM capability for photonic cross-connects. When an edge node signals an LSP, relying on the Interface Switching Capability Descriptor provided by the TE-LSDB, it can find out whether the interface the signaling comes in has different switching capability from the outgoing interface. If so, it knows it may be at the boundary of two levels of LSP. For example, an edge node may have the Interface Switching Capability Descriptor of its interfaces like:

Descriptor for Interface 1:

Interface Switching Capability = TDM

Encoding = SONET

Max Bandwidth[0] = 10 Gbps, for priority 0

Descriptor for Interface 2:

Interface Switching Capability = FSC (Fiber Switch Capable)

Encoding = Ethernet 802.3

Max Bandwidth[0] = 100 Gbps, for priority 0

When the signaling message comes in from interface 1 and the outgoing interface for it will be interface 2, the edge node understands that a hierarchical LSP will be established (see the example in Figure 3.13). The low-order LSP is tunneled through the high-order LSP, and multiple low-order LSPs can be aggregated into the high-order LSP.

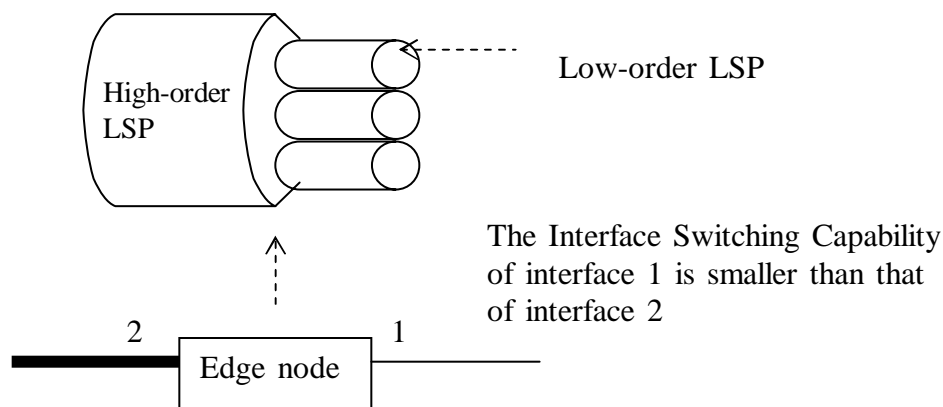


Figure 3.13: the edge node knows if a hierarchical LSP will be established

Here we illustrate how the hierarchical LSP set up is done using RSVP-TE with extensions to support GMPLS. Lower-order LSPs trigger the set-up of a higher-order LSP. Nodes at the border of two different networks in terms of multiplexing capabilities are responsible for establishing higher-order LSPs and aggregate lower-order LSPs. Figure 3.14 shows an example. Packet-Switch Capable (e.g., IP packets) LSR 1 and 2 are connected by a 500 Mb/s Ethernet link, so are LSR 7 and 8. SONET switches and LSRs are connected by OC-12 links; SONET switches and PXC are connected by OC-192 links; PXC are connected by optical fibers. Note that PXC 4 and 5 may not be connected directly, e.g., there are other PXC between the two. Let us assume that the edge PXC has the capability to convert electrical signals to optical signals. They have interfaces that can provide SONET signals and interfaces that can provide WDM capability. An LSP (LSP 1) is going to be established from LSR 1 to LSR 8, which requires 500 Mb/s bandwidth.

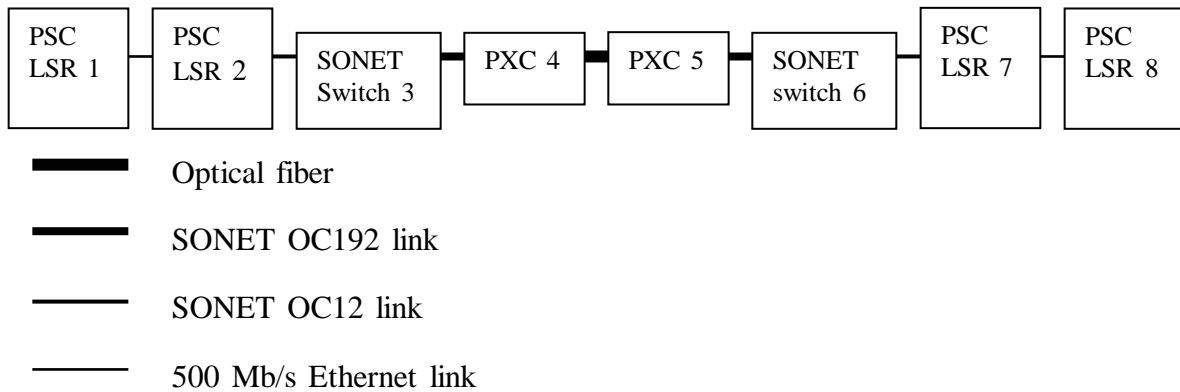


Figure 3.14: a hierarchical LSP is established between LSR1 and LSR8

We assume that all links have enough bandwidth for the LSPs to be established, and that there is no existing LSP between the different nodes. The GMPLS signaling using RSVP-TE starts from LSR1 (see the following figure).

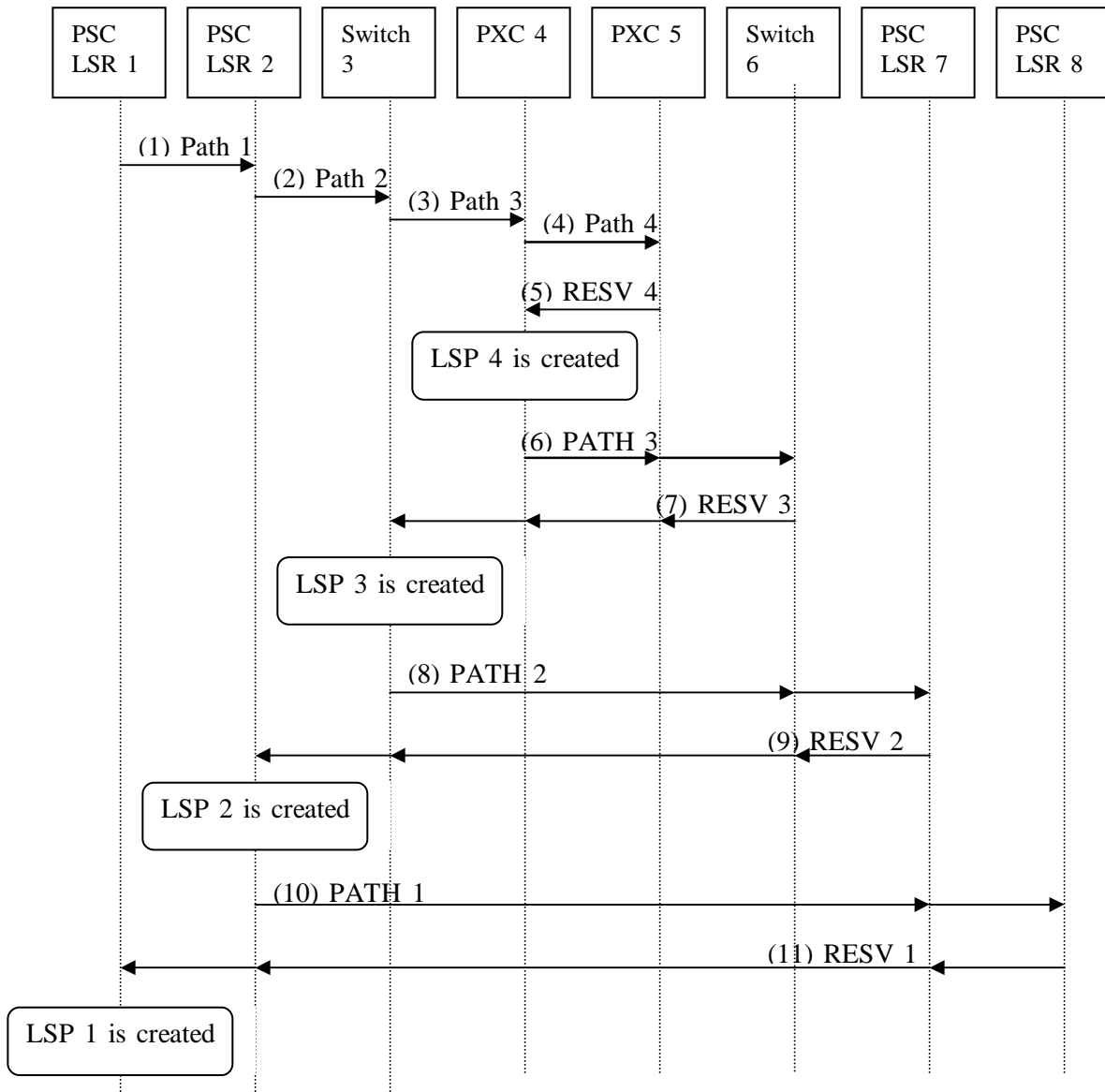


Figure 3.15: the time-sequence of establishing a hierarchical LSP

(1) The RSVP-TE Path message (Path 1) generated by Router 1 arrives at Router 2. This is the Path message for LSP 1, so let us call it Path 1. Based on the link information from the TE Link State Database, Router 2 knows that the path must cross links that are different (e.g., different types of interface, bigger multiplexing capacities). So Router 2 is triggered to establish a new path LSP2 that will be terminated on Router 7. This represents the next-higher LSP through which the LSP from Router 1 to Router 8 will be multiplexed. Router 2 generates another RSVP-TE Path message (Path 2 for LSP 2) destined to Router 7.



(2) Path 2 arrives at SONET Switch 3. Again, Switch 3 finds out that the LSP must cross different links. Switch 3 is going to establish LSP 3, and it generates RSVP-TE Path message destined to Switch 6 (Path 3).

(3) Path 3 arrives at PXC 4, which triggers PXC 4 to establish LSP 4. So PXC 4 generates a Path message destined to PXC 5 (Path 4).

(4) Path 4 arrives at PXC 5.

(5) PXC 5 responds with a RSVP-TE RESV message. Let us call this RESV message Resv 4. Resv 4 arrives at PXC 4, and the LSP 4 is created. LSP 4 is a TE link. It will be advertised by the routing protocol that runs at this level (e.g., the network constructed by the PXC's) as a lambda-switch-capable link. This LSP is a FA-LSP. The capacity of this TE link in the advertisement is the difference between its maximum capacity (e.g., a number of lambdas) and the share (e.g., one lambda) that has been allocated for the OC-192 bandwidth.

(6) Then PXC 4 continues signaling for LSP 3. The PATH message Path 3 goes on.

(7) Path 3 arrives at Switch 6, and Switch 6 responds with a RSVP-TE RESV message. Let us call it RESV 3.

(8) RESV 3 arrives at Switch 3. LSP 3 is created. LSP 3 is a TE link. It will be advertised by the routing protocol that runs at this level (e.g., the network constructed by OC-192 switches) as a TDM-switch-capable link. This LSP is a FA-LSP. The capacity of this TE link in the advertisement is the difference between its maximum capacity (e.g., OC-192 bandwidth) and the share (e.g., OC-12 bandwidth) that has been allocated for the LSP 2 being established. Then the LSP 2 set up procedure continues, and Path 2 goes on to Router 7.

(9) Router 7 responds with a RSVP-TE RESV message (RESV 2).

(10) RESV 2 arrives at Router 2 and LSP 2 is created. LSP 2 is a TE link. It will be advertised by the router protocol that runs at this level (e.g., the network constructed by OC-12 switches) as a TDM-switch-capable link. This LSP is a FA-LSP. The capacity of this link in the advertisement is the difference between its maximum capacity (e.g., OC-12 bandwidth) and the share (e.g., 500 Mb/s) that has been allocated for the LSP 1 being established. Then the LSP 1 set up procedures continues and Path 1 goes on to Router 9.

The hierarchical LSP established in the above example is illustrated in Figure 3.16.

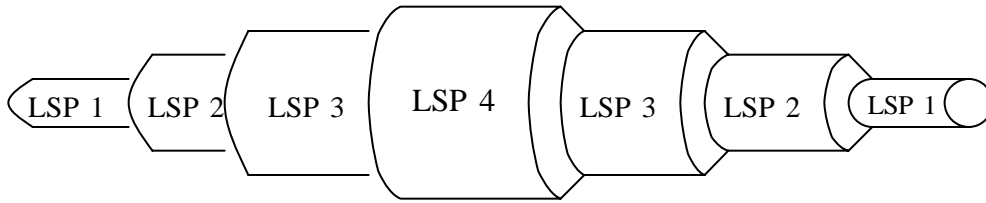


Figure 3.16: the hierarchical LSP in the example

If there is an existing FA-LSP that can satisfy the LSP being established, e.g., its unreserved bandwidth is bigger than what the LSP being established needs, then the edge node is responsible for tunneling the low-order LSP onto the existing high-order FA-LSP. In the above example, if LSP 4 has already been established (between PXC 4 and 5) when the Path message for LSP 3 (Path 3) arrives, then LSP 4 is treated as a single link and the Path 3 message will take PXC 5 as its destination, which is the ending node of LSP 4.

If the LSP being established is an explicit-routed LSP (ER-LSP), the RSVP-TE Path message carries an Explicit-Routed Object (ERO). A node receiving this message determines if it is the edge node at the boundary of two LSPs. If it is not, the conventional signaling goes on. If it is, it must determine which node is the ending node of the high-order LSP (the other edge). Then it must extract from the ERO the subsequence of hops from itself to the edge of the network. Let us call this subsequence of hops S1.

An example is shown in Figure 3.17. Node 1, 2 and 3 are part of an optical network. The RSVP-TE Path message carrying the ERO arrives at Node 1. Node 1 checks the nodes in the ERO one by one. From the routing database, it finds out the first 3 nodes starting from the beginning of the ERO are in the same network.

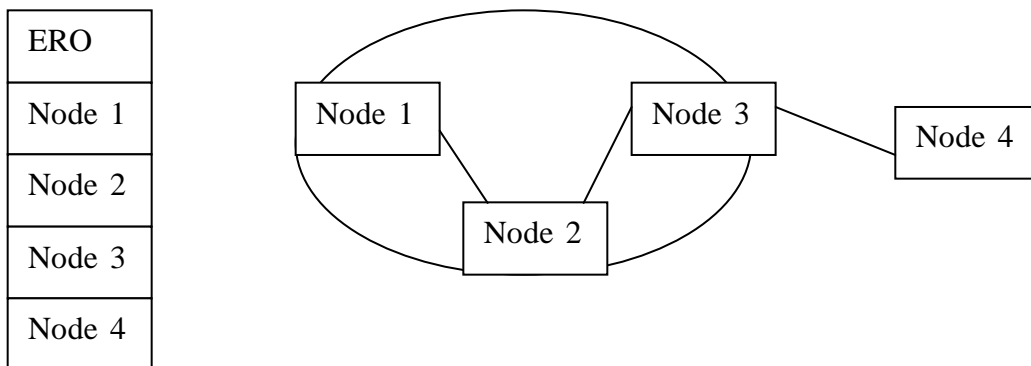


Figure 3.17: the ERO drives RSVP-TE to establish a hierarchical ER-LSP

Then the edge node checks the TE LSDB to see if there is an existing FA-LSP whose hops exactly match S1. If there is, it further checks if the properties of that FA-LSP can meet the requirements of the LSP being established, e.g., the unreserved bandwidth can satisfy the LSP being established. If so, the node replaces the hops S1 in the ERO with the end node of the FA-LSP. In the above example, Let us assume that there is a FA-LSP constructed by Node 1, 2 and 3. Node 1 replaces Node 1, 2 and 3 with Node 3 in the ERO. Then the destination address of the Path message is set to Node 3, and sent out by Node 1. We can see that the FA-LSP is treated as one link. After that, the TE properties of the FA-LSP are adjusted, e.g., the unreserved bandwidth is the difference between the previous unreserved bandwidth and the requirement of the LSP being established. They are advertised by the routing protocol in the current routing domain, e.g., by OSPF TE-LSA.

If there is no existing FA-LSP or the existing FA-LSPs do not satisfy the requirement of the LSP being established, then the edge node will signal a new high-order LSP, which will tunnel the low-order LSP. And it would be advertised as a FA-LSP.

The unreserved bandwidth of the FA-LSP is the difference between the maximum reservable bandwidth and what all the multiplexed low-order LSPs request.

The FA-LSP should be torn down if no tunneled LSP is there. There are a number of ways to trigger the FA-LSP tear-down. For example, if the maximum reservable bandwidth is as same as the unreserved bandwidth, then it means the FA-LSP is not being used, and it should be torn down.

#### **4.3.1.1 Signaling Support for Fault Notification**

The Notification mechanism in the signaling protocol RSVP-TE [26] is dedicated to support the fault notification in GMPLS recovery.

Fault notification is to notify the nodes of the failure in the path that are responsible for recovery. RSVP-TE defines a rapid fault notification mechanism to convey the information. The Notification mechanism includes the Notify Request object and the Notify message.

The Notify Request object contains the IP address of the node that should be notified of the failure, which is named *Notify Node Address*. This address can be configured, or automatically determined by the protection mechanism. For example, in the 1+1 protection mechanism, the LSP initiator node is responsible for switching the traffic to the backup LSP when the failure occurs, so the *Notify Node Address* should be that node. The LSP initiator node may be responsible for attaching this request object in the RSVP-TE Path message. The receiver of such a Path message (transit nodes) should also attach this object to the outgoing Path message. So the request is propagated. The terminator node of the LSP may respond with a Resv message which also carries the Notify Request object for a bi-directional LSP. So the notification may be required in both directions. A

node receiving the message records the *Notify Node Address* in the protocol soft state (for the RSVP soft state, see the RSVP introduction in the previous sections).

The Notify message provides a mechanism to inform non-adjacent nodes of LSP related events. It is different from the RSVP error message. The RSVP error message must be forwarded one by one along the nodes of the LSP, which is too slow for fault notification and not necessary. Notify message can be “targeted” to a particular node, e.g., the traffic-switch-over trigger node. By “targeted” it means the destination address of the IP packet carrying the Notify message is set to the IP address of the target node, which is specified by the *Notify Node Address* received from the Notify Request object. So it does not need to travel along the hops of the original LSP. Because after a failure in the network, the network topology likely has changed and there is another path that is optimal for the Notify message (see Figure 3.18). Nodes receiving a Notify message, which is not the destination of the message, must forward the message, unmodified, to the target.

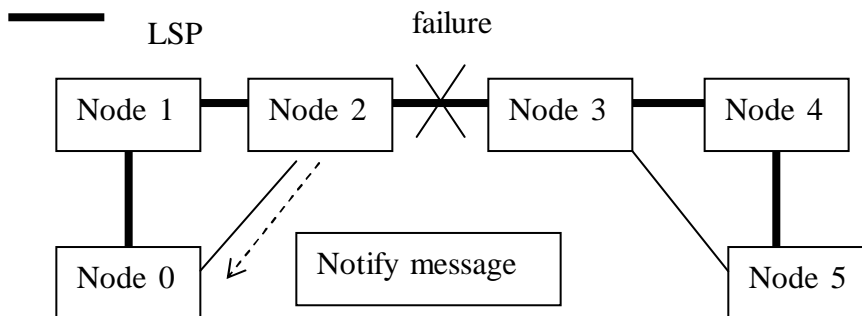


Figure 3.18: the Notify message is sent to the targeted node directly

The Notify message contains an `ERROR_SPEC` object, which specifies the IP address of the node that detects the failure or the link that has failed. Optionally it may carry other RSVP-TE objects that describe the LSP, e.g., the `LSP_SESSION` object. Notify messages are normally generated only after a Notify Request object has been received.

It is not necessary for the local recovery to use such a notification mechanism. But other mechanisms need it, for example, the end-to-end LSP protection. Section 5 will specify which recovery mechanism needs it and when.

#### 4.3.2 GMPLS signaling: CR-LDP with extensions

RSVP-TE, as a label distribution protocol, was built on the existing control protocol RSVP (RFC2205) [40]. Label Distribution Protocol (LDP) [41] was originally designed for label distribution.

LDP also uses the TLV structure to encode messages, which allows for future extensions. At first, LDP discovers its peers by multicasting an LDP Hello message onto the network. The nodes running LDP that receive the message are triggered to establish an LDP

session with each other. After the session is successfully created, they become LDP peers, and the session is maintained. Then the LDP peers can exchange label distribution messages. If there is any error during label distribution, the LDP Notification messages are used to provide error information, which could tear down the LDP session between LDP nodes. LDP uses TCP as the reliable transport mechanism to deliver all messages except the LDP Hello message, which uses UDP.

LDP has been extended to support Traffic Engineering, which is named *Constraint-Based LSP Setup using LDP* (CR-LDP) [42]. CR-LDP defines a new set of TLV structures to support explicit routed signaling, traffic parameters, LSP set-up/holding priority, etc. It also defines a means for resource reservation.

The constraint-based route TLV structure contains a sequence of IP prefixes or a sequence of Autonomous Systems. The contents of the constraint-based route TLV are computed by CSPF, which tells the routing mechanism where to forward the CR-LDP messages.

The LSP signaled by CR-LDP is initiated by the head node of the LSP. How it works is illustrated as below.

Router 1 is going to establish an explicit-routed LSP (R1, R2, R3, R4, R5). R1 constructs the constraint-based route TLV to have the sequence of nodes (R1, R2, R3, R4, R5). Each node can be represented by an IP address. Then R1 sends out the CR-LDP Label Request message carrying the constraint-based route TLV. Before the message is sent, R1 checks the top of the TLV, and it finds out that the next hop is R2. R1 removes itself from the TLV and sends the message to R2. The Label Request message may carry the Traffic Parameter TLV, which specifies the traffic parameters to be sent. If so, R1 reserves the resource before the message is sent out. R2 receiving the message also checks the top of the TLV, and it finds out R3 is the next hop. R2 removes the address of R2 from the TLV and sends out the Label Request message. It may also reserves the resource for the LSP if the Traffic Parameter TLV is carried. R3, R4, and R5 follow the same algorithm as R2 does. When R5 receives the message, the TLV only has one address, which is R5 itself. Along the message path, the LDP protocol state should be created.

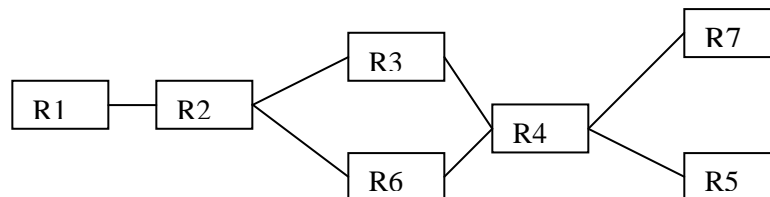


Figure 3.3.2.1: ER-LSP from R1 to R5

R5, as the ending node of the LSP, programs the label forwarding table, reserves the resource if needed, and responds with a CR-LDP Label Mapping message, which carries a

Label TLV. The Label TLV contains the label that the downstream node wants the upstream node to use. The protocol state on the node tells R5 to send the Label Mapping message to R4. R4, R3 and R2 do the same thing as R5 does. R1, as the head node of the LSP, does not need to allocate label any more, but simply receives the label and programs the label forwarding table.

The head node of an LSP transmits a Label Release message to a peer when it is no longer needs a label previously received from that peer. This takes place when the LSP is torn down (see the following figure).

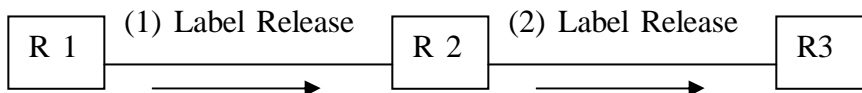


Figure 3.19: LSP (R1, R2, R3) is torn down by Label Release message

Unlike the RSVP-TE, CR-LDP is not a soft state protocol. By this it means the LSP created by CR-LDP does not need the signaling refresh periodically. The LSP must be torn down explicitly.

CR-LDP is also being extended to support GMPLS [43]. The information that is needed to support generalized label in RSVP-TE is also needed for CR-LDP. For example, the label format in the Label TLV is also generalized to support different types of “label”, e.g., the port number, wavelength, etc.

CR-LDP is also required to support bi-directional LSP set up. The idea is to add the Upstream Label TLV in the Label Request message (see Figure 3.20).

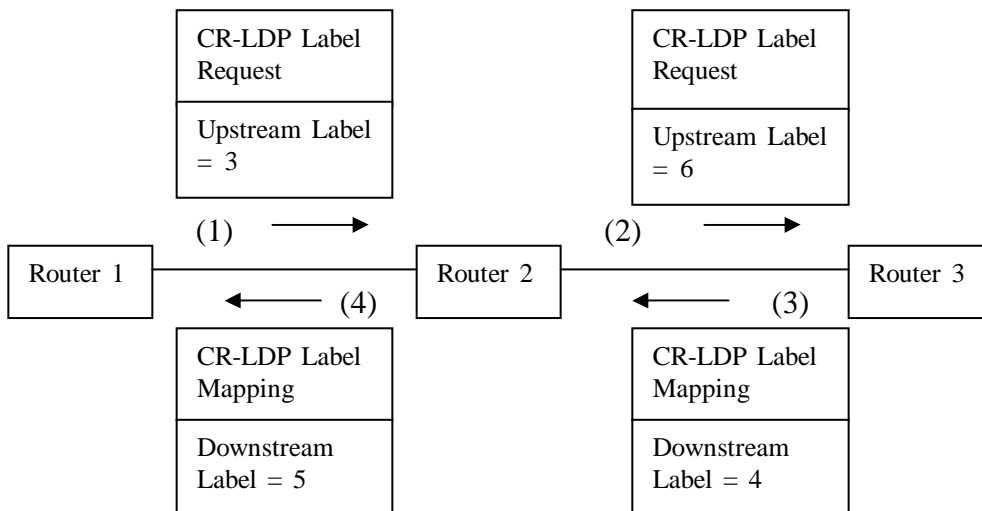


Figure 3.20: CR-LDP signals a bi-directional LSP

To support explicitly routed LSP in the context of GMPLS, just the IP address or the identifier of an autonomous system may not be adequate. RSVP-TE defines the Label object as a sub-object in the ERO, and CR-LDP defines the Explicit Label Control TLV as a sub-TLV following the IP address or the Autonomous System ID in the constraint-based route TLV.

In order to improve network survivability, the protection information is considered in GMPLS signaling. Like RSVP-TE, CR-LDP defines the Protection TLV, which includes:

- (1) link protection type;
- (2) indication of whether the path is primary or backup.

Can *CR-LDP with extensions* [25] do whatever *RSVP-TE with extensions* [26] can do so as to support GMPLS signaling? No, as this report is being written. The RSVP-TE [26] has got the fault notification mechanism (see Section 4.3.1.1) to notify a responsible node when a link/node failure occurs. But CR-LDP [25] does not have a similar mechanism yet. CR-LDP [25] has its own Notification message, but it does not provide the same function as the one does in RSVP-TE.

From now on in this report, RSVP-TE is used to illustrate the GMPLS signaling support for LSP protection/restoration.

#### **4.4 The Hello Protocol**

In fact, there is no protocol called Hello. OSPF, RSVP-TE, LMP and other protocols define a software method to detect failures, which is the Hello messaging. The idea of the Hello messaging is simple. Two nodes exchange a short message named Hello periodically. The interval can be configured, e.g., the recommended interval for OSPF Hello is 5 ms (see RFC2328). If a number of messages are missed, e.g., 4, then the node can determine that the other node is down or the link between the two is broken.

Although many protocols provide this method to detect failures, using software to detect a failure is very slow and usually does not meet real-time application requirements. Furthermore, it is difficult for the software detection to deal with the shaking problem. A node does not receive OSPF Hello messages from its neighbor for several times, and it determines its neighbor is down. But just after that it can receive Hello again due to the unstable situation in the network. The problem keeps repeating like that for a while, which is called *shaking*.

However, the software fault detection is still useful in some situations. An example is the Ethernet, where nodes are connected by a bus (multiple access media). A node can detect its peer's fault by the Hello messaging.

## 5. The Recovery Mechanism in GMPLS

There are two recovery mechanisms: protection and restoration.

**Protection:** A dedicated protection path is established for a connection, and the connection switches from the working (primary) path to the protection (backup) path when a failure occurs on the primary path.

**Restoration:** The establishment of a backup path does not occur until a failure occurs in the primary path. Then the traffic is switched to the backup path. Such a mechanism is called restoration. But the backup path can be selected (calculated) in advance.

Restoration and protection are different mechanisms. They operate on different time scales; protection requires redundancy of resources, while restoration relies on dynamic resource reservation - hence restoration takes more time [44].

Protection/Restoration can be classified into the following categories according to the recovery scope (see [12] and [45]):

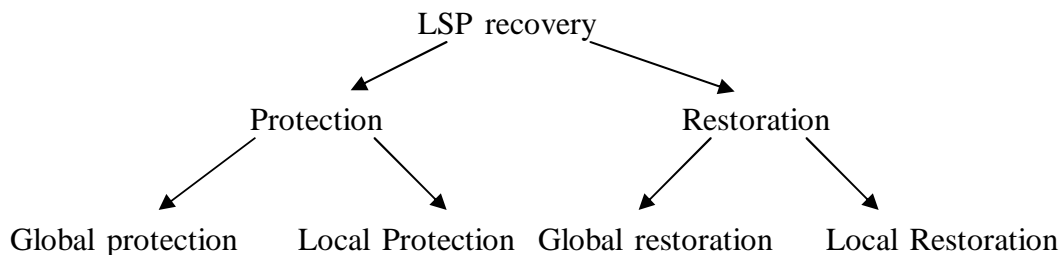


Figure 4.1: LSP recovery

The objective of local recovery is to protect against a link or neighbor node fault and to minimize the amount of time required for fault notification. The local recovery includes link recovery and node recovery. Local recovery is initiated by the immediate upstream node of the faulty link or node, which may be a transit node or the source node of an LSP.

The objective of global recovery is to protect against any link or node fault on an LSP or on a segment of an LSP except for the source or the destination node. Global recovery is also called end-to-end path recovery, because only the source or destination node initiates the recovery process.

### 5.1 Protection Mechanisms

The protection mechanisms are described in the following. The ideas can be applied on paths as well as links. These mechanisms can be used in any network that may have different switching technologies at any level of the GMPLS hierarchy, for instance, ATM networks, IP networks, optical (e.g., OXC) network, etc.



- 1+1 protection

Two disjoint paths have been established and both of them have resources allocated. By “disjoint”, we mean none of the links or nodes constructing these paths are overlapped (except the starting node and the terminating node of these paths). The same user data is transmitted simultaneously over the two paths, and the receiver can pick the best signal from one of these two paths. An example can be seen in Figure 4.2. In the example, Path 1 and 2 provide a 1+1 protection for the data transport from Node 1 to Node 5. If Path 1 is broken, for instance, then the receiver at Node 5 can pick the signals from Path 2.

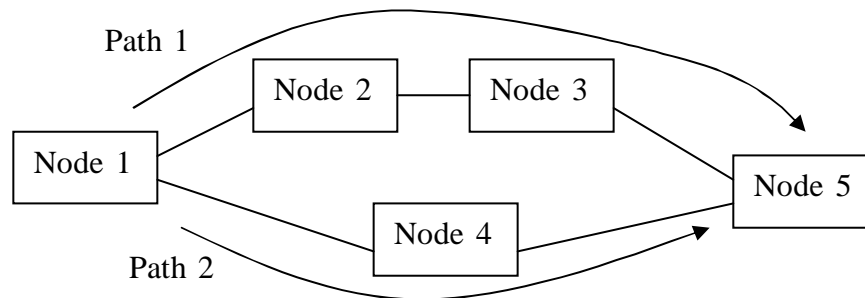


Figure 4.2: 1+1 path protection

The benefit of 1+1 protection is short recovery time and the lost data can be very small. But it requires two pre-established paths, double resources and the traffic is copied and sent over both paths. It is expensive.

- M : N protection

There are M pre-established backup paths that protect N primary paths. But user traffic is not transported by any of the backup path until a failure occurs. When one of the primary paths fails, the nodes connecting to the faulty link or the faulty node notify the end-nodes of the path (source and destination nodes). Then the end-nodes allocate the resource required by the traffic traveling that primary path on one of the backup paths. In the end, the traffic is switched over. Note that the backup paths can protect any of the primary paths. An example can be seen in Figure 4.3. In the example, 2 backup paths (Path 1 and 2) are protecting 2 primary paths (Path 3 and 4). If Path 4 is broken (for instance, the link between Node 7 and 8 is broken), then Node 7 notifies Node 1 to do the protection switch (and maybe Node 8 notifies Node 5 as well - depending on how the signaling protocol works). Node 1 allocates the resource on Path 1, which is required by the traffic traveling on Path 4, and switches the traffic from Path 4 onto Path 1.

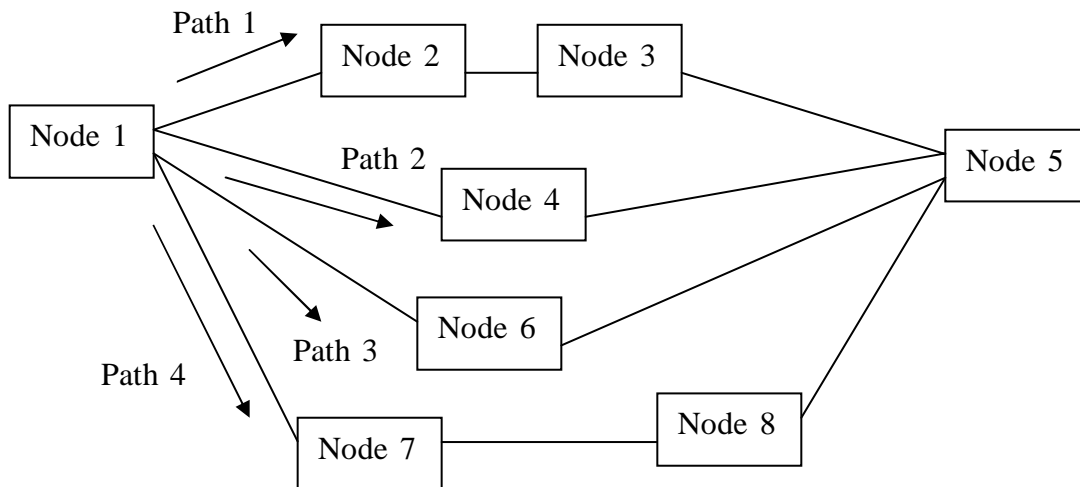


Figure 4.3: 2:2 path protection

It is not recommended that the links constructing different primary paths belong to the same Shared Risk Link Group. For example, both primary path (L1, L2, L3) and (L4, L2, L5) have the same link L2, and they would better not share the same backup path. Otherwise, if L2 goes down, it is possible that one of the primary paths would not have any protection.

- 1 : N protection

It is a special case of M : N protection - only one pre-established backup path provides protection for N primary paths. Let us look at Figure 4.3 again, and assume that Path 1 is protecting other paths. If the link between Node 7 and Node 8 is broken, Node 7 notifies Node 1 (and maybe Node 8 notifies Node 5 as well) to do the protection switch. Then Node 1 allocates resource required on Path 1, and switches the traffic onto Path 1.

If there are two primary paths that are broken simultaneously, then a policy is used to decide which one will get protected. One of the policies is taking priorities. For example, Path 3 and 4 are broken, if the traffic traveling on Path 3 is deemed to have higher priority than the traffic traveling on Path 4, then Path 3 will get protected by Path 1. Another simple policy can be First-Come-First-Service.

The links constructing the primary path should not belong to the same Shared Risk Link Group. Otherwise, if the link constructing both of the paths is cut, then one of the primary paths does not have any protection.

- 1 : 1 protection

It is also a special case of M : N protection - one dedicated backup path is pre-established for one primary path. For optimization, the resource may be pre-allocated if it is known in advance. But the user traffic is not inserted onto the backup path. So the resource pre-allocated on the backup path may be used by other LSPs that have lower

priorities. When the primary path fails, the signaling protocol notifies the end-nodes of the primary path. Then the traffic is switched over from the primary path and the LSPs that are using the resource of the backup path are preempted.

### Summary of Protection Mechanisms

When a failure occurs, the nodes involved in the recovery need not notify the end-nodes of the route (path) in the 1+1 protection mechanism; but in the M:N, 1:1 and 1:N protection mechanisms, the nodes neighboring the failure must notify the end nodes so that the end-nodes will switch the traffic. So the 1+1 protection mechanism provides fast recovery because it does not need fault notification time. However, the other mechanisms utilize the resources more efficiently.

#### 5.1.1 Local Protection

Local protection includes link protection and node protection.

##### 5.1.1.1 Link Protection

Link protection switches the traffic from the primary link to a backup link between the same nodes when link failure occurs. It occurs between two adjacent nodes and only these two nodes are involved in the whole process.

As we specified in the sub-section *Enhancements in MPLS Signaling to Support GMPLS* (see Section 2.4.3), the requested link protection type is carried by the signaling protocol when an LSP is set up. The node that receives such a request must check the outgoing interface to see whether the request can be satisfied. If the link protection request is not satisfied, then the signaling for the LSP establishment cannot continue.

For RSVP-TE signaling, the Path message carries the link protection type for the LSP. The protection object of RSVP-TE signals the link protection type and the role of the LSP (see Figure 4.4). The S bit signals the role of the LSP being established and the “link flags” signal which link protection type is desired. If bit S is set, it means the LSP is a secondary (backup) one; otherwise, it is a primary LSP. The link protection flag contains one of the codes specified in Section 4.1.3.3.

S	reserved	Link protection flags
---	----------	-----------------------

Figure 4.4: the protection object in RSVP-TE

An example is shown in Figure 4.5. In the example, the Path message carries a protection object to establish an LSP. The protection object signals the link protection type is “Dedicated 1+1” and the LSP being established is a primary one. The Path message arrives at the node. The node must check if there is a link connecting the next hop, which has link protection capability “Dedicated 1+1”. Because the link protection type is distributed by the routing protocol, for example, the node checks the link state database maintained by the routing protocol. According to the definition of protection type

Dedicated 1+1, we know that the protecting link must not be in the same Shared Risk Link Group (SRLG) as the primary link. If there is such a link, signaling continues.

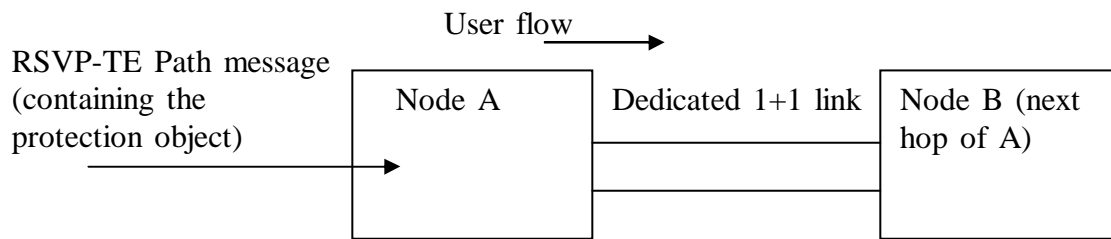


Figure 4.5: the link protection type must be honored to continue signaling

When the RSVP-TE Resv message arrives at Node A, it reserves the resource (e.g., bandwidth) on both of the links. After the LSP is created, the node (in this example, Node A) will copy the traffic and insert it into both links. The receiver selects the healthy traffic from any of the links. For example (see Figure 4.6), after initialization, the receiver takes the traffic from the primary link. When the primary link fails, LMP Fault Management (see the sub-section about LMP) is used to localize the failure. For example, all the nodes following Node B can detect loss of light if the nodes are in the optical network. LMP tells Node A and B that the faulty link is between them. Node B simply selects the traffic from the backup link.

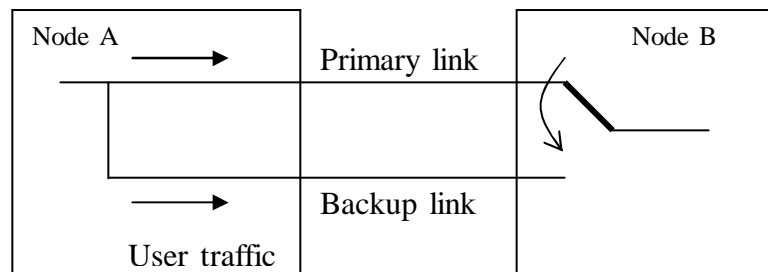


Figure 4.6: Dedicated 1+1 link protection

If the link protection type is shared, e.g., 1:N (or 1:1), then the Resv message also reserves the resource on the backup links. But the backup links will not transport traffic. And the resource reserved for the backup links can be used by other LSPs that have lower priorities. The reason is that these lower-priority LSPs will be preempted when the primary links fail, and the traffic is switched over.

With link layer protection, the LSP may stay there even though there is a link failure, and LSP recovery mechanism is even unaware of the problem. The failure will be reported by alarms signaled by the network management in the nodes connecting the faulty link. In the example, the alarms will be displayed on Node A and B.

## Summary of Link Protection

Because the point that initiates the recovery is close to the failure, there is no need to have fault notification - it provides fast recovery. Only the nodes connecting the faulty link are involved in the recovery. And it does not require any changes in the GMPLS LSP.

But the protection ability is limited. If the entire LSP requires link layer protection, it is expensive and the control becomes a big overhead, because every node along the whole LSP needs to keep monitoring links.

Therefore, usually link layer protection is only used in an area that is deemed to be unreliable.

### 5.1.1.2 Node Protection

In fact, there is no protection mechanism in the GMPLS LSP level that is dedicated to locally provide single node failure protection. If some nodes in a network are deemed to be unreliable, then the path computation should compute a path that will work around those nodes. On the other hand, global path protection and restoration can recover the traffic affected by node failure. These mechanisms will be introduced later.

### 5.1.2 Global Protection

From the previous sub-section, we can see that link layer protection provides the link protection under the GMPLS LSP layer.

With the end-to-end path protection (global protection), multiple disjoint hierarchical LSPs are pre-computed and established between the initiator and the terminator nodes of a client LSP. Dedicated resources are allocated for these LSPs. So the nodes and links of the entire primary hierarchical LSP are protected except for the initiator and terminator nodes. In order to avoid the contention of multiple layer protection mechanisms, the LSPs may require “unprotected” Link Protection Type during signaling. Thus the protection is only built on the MPLS-based layer and contention will not occur. When a failure occurs, the nodes that detect the failure notify the end nodes (the initiator and terminator nodes). The end nodes initiate switching the traffic to the alternate path.

The illustration is shown in Figure 4.7. The logical view of the 1:1 LSP protection is shown in the figure. The LSP (Node 1, Node 3, Node 5, Node 7) is the primary one; LSP (Node1, Node2, Node4, Node6, Node8, Node7) is the backup. Both may be hierarchical LSPs, e.g., the link (Node3, Node5) is a FA-LSP (TE link), so is link (Node4, Node6). Traffic is sent along the primary LSP. If a failure occurs, the nodes that detect the failure notify the end nodes: Node 1 and 7, then the end nodes will switch the traffic to the backup LSP.

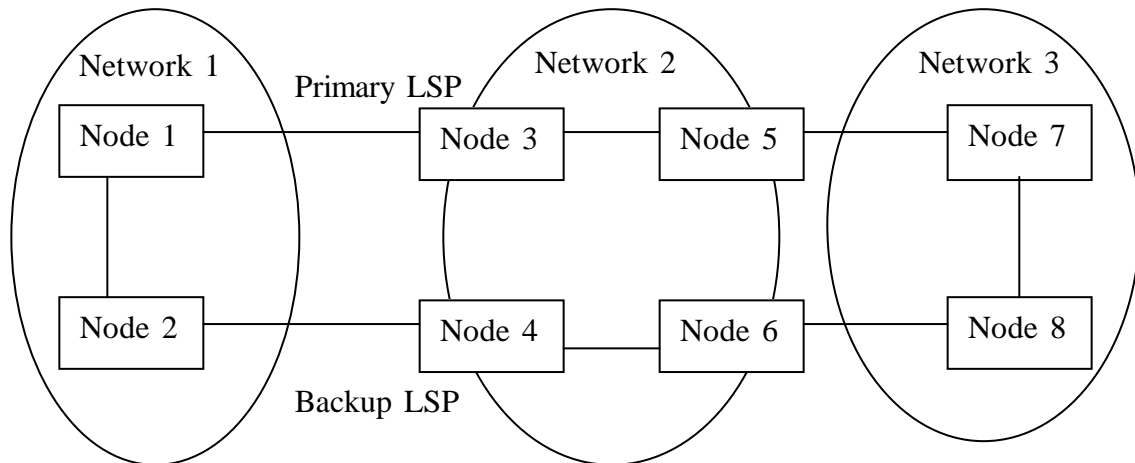


Figure 4.7: The logical view of the 1:1 LSP protection.

### Summary of Global Protection

Global protection can provide a fast protection mechanism against any link or node fault on an LSP with the exception of the failure occurring at the initiator and terminating node (end nodes) of an LSP. Usually, the end nodes are far away from the failure, and need to be notified by the node that detects the failure, which takes time. Also, it is expensive because the backup path is pre-computed and pre-established. The resource is pre-allocated as well, but it may be used by low priority traffics.

### 5.2 Restoration Mechanisms

Restoration is implemented by rerouting. Some papers even use the term rerouting [31]. Rerouting is referred to as establishing new paths (global restoration) or path segments (local repair) on demand for restoring traffic after a failure occurs.

Rerouting follows the “make-before-break” principle. The “make-before-break” means the original path is used while the new path is set up, then the node performing the reroute switches the traffic to the new path and the original path is torn down.

#### 5.2.1 Local Restoration

Local restoration includes link restoration and node restoration. When a link failure occurs between two adjacent nodes, with link restoration, the upstream node switches the traffic on an alternate route in which there are additional intermediate nodes. In the case of node failure, the immediate upstream node of the faulty node initiates an alternate route, which bypasses the faulty node. Then traffic is switched over to the alternate route. Such rerouting also provides the local restoration for node failure.

Upon detecting a failure, paths or path segments to bypass the failure are established using signaling. The idea is shown in Figure 4.8. Assuming that the path is (Node 0, Node 1, Node 2, Node 3, Node 4). If Node 2 is down, Node 1 creates a path segment which bypasses the faulty node – (Node 1, Node 5, Node 3). The new path segment goes

through another interface of Node 1 and arrives at Node 3 through another interface. For example, using RSVP-TE, because the message carries an identification (e.g., the Session object and the Sender Template object in RSVP-TE) for each LSP, the Path message can re-create the protocol state in Node 3 and re-program the label forwarding table. The original path segment will be torn down eventually.

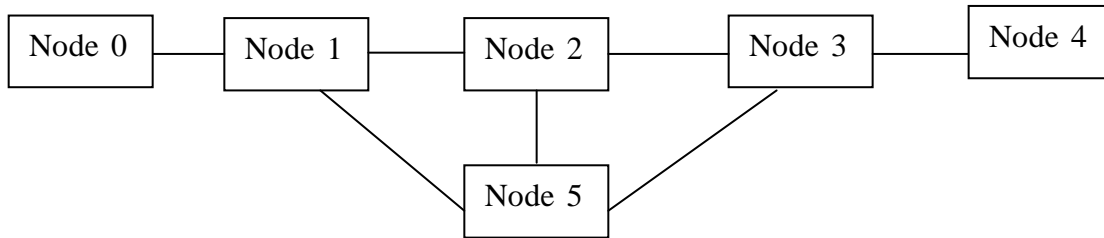


Figure 4.8: reroute

According to the position of the faulty node in the LSP, there are three cases.

**Case 1:** The failure does not occur at the end node of the hierarchical LSP.

In this case, there is no difference between link and node restoration from the rerouting point of view. An example is shown in Figure 4.9. In this example, if OXC 3 fails or the link between OXC 3 and OXC 4 fails, Case 1 occurs.

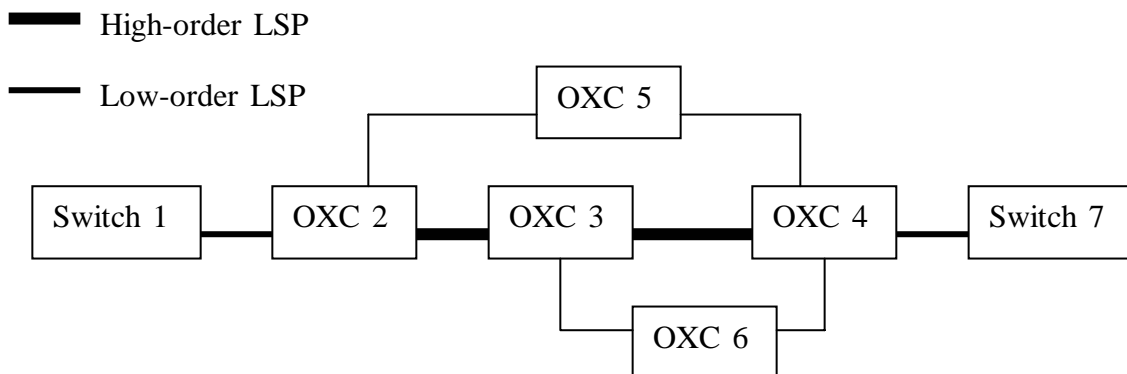


Figure 4.9: a situation where reroute Case 1 applies

The recovery steps are:

- (1) The failure detection mechanism detects the failure.
- (2) The fault localization mechanism localizes the failure. Meanwhile, the node that is the immediate upstream node of the failure knows about the failure.
- (3) The node initiates the process to establish a new path or path segment that bypasses the failure.
- (4) And the node switches the traffic to the alternate path.

As in all the recovery mechanisms, the failure detection usually is done by hardware at the physical layer (or link layer). After that, the fault localization mechanism is triggered,

which can find out where the failure is. For example, LMP fault management (see the Section 4.2) can localize a link failure. The fault localization mechanism does not stop running until the node that is the immediate upstream node of the failure finds out the failure. So there is no need to have explicit fault notification. For a faulty node, the immediate upstream node of the faulty node detects the problem. Then signaling is used to create a reroute.

Using RSVP-TE, the reroute initiator node sends out the Path refresh message, which will consult the routing component for a feasible route. In the example in Figure 4.9, if OXC 3 is down, OXC 2 detects the failure, e.g., by OSPF Hello messaging or other means, and sends out the Path refresh message, which can travel to OXC 5 to get to OXC 4. OXC 4 responds with a Resv refresh message, and the LSP between OXC 2 and OXC 4 is fixed. If the link between OXC 3 and OXC 4 is broken, LMP fault management can localize the failure. OXC 3 sends out the Path refresh message, which can travel to OXC 6 to get to OXC 4 and repairs the FA-LSP.

**Case 2:** The terminator node of a FA-LSP fails. An example is shown in Figure 4.10. In this example, if OXC 4 fails, Case 2 occurs.

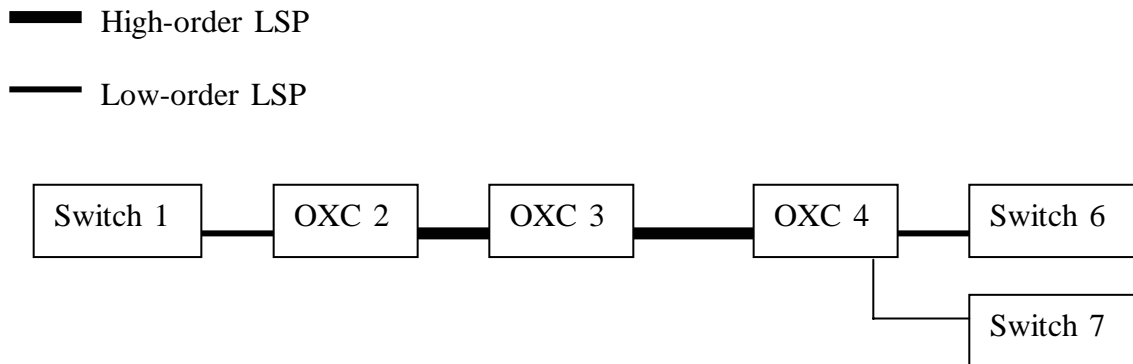


Figure 4.10: the situation where reroute Case 2 applies

Note that the high-order LSP between OXC 2 and 4 is an FA-LSP, which tunnels low-order LSPs. As the OXC 4 is the terminator node of the FA-LSP, it is impossible to rebuild this FA-LSP so that it meets the requirements of the tunneled LSPs. For example, the tunneled low-order LSPs go to multiple different destinations after the FA-LSP, like, Switch 6, Switch 7, etc. The FA-LSP does not know the information, so there is no sense for OXC 2 to reroute FA-LSP and there is no need to repair the FA-LSP.

However, local recovery can still be useful. The OXC 2 is the node that tunnels (aggregates) a number of low-order LSPs. If the FA-LSP is broken, the OXC 2 can trigger all the tunneled LSPs to reroute individually. For example, a low-order LSP, which was tunneled by the FA-LSP at OXC 2, can re-establish a path segment that bypasses the failure and reaches the desired destination, e.g., Switch 6. Let us see how it works.



When OXC 4 fails, it is as if the “link” *FA-LSP* failed. Because OXC 2 and OXC 4 have the “Forwarding Adjacency” (FA), OXC 2 should be notified according to the link restoration mechanism. The Notification mechanism of RSVP-TE is useful here.

OXC 2, as the initiator node of the FA-LSP, can attach the Notify Request object to the Path message when the FA-LSP is established, and the target node address in the object is OXC 2 itself. When OXC 4 fails, fault detection, e.g., OSPF Hello messaging, enables OXC 3 to detect the neighbor failure. Then OXC 3 notifies OXC 2.

Another way is by administration. During signaling, OXC 3 knows that it is the penultimate node of the FA-LSP, e.g., routing tells OXC 3 that it is directly connected to OXC 4. Let us assume that we have such an administration policy that the penultimate node of the FA-LSP must notify the initiator node of the LSP. The Notify message destination address can be configured. In the example, OXC 3 can send out the RSVP-TE Notify message targeted to the initiator node - OXC 2 in this example.

After the initiator node of the FA-LSP is notified, it tells all the tunneled low-order LSPs to re-establish the LSP segment (e.g., maybe another hierarchical LSP) that bypasses the fault.

**Case 3:** The initiator node of a FA-LSP fails.

If the initiator node of a client LSP fails, then there is no general LSP protection/restoration mechanism.

If the initiator node of a FA-LSP fails, then the immediate upstream node of the faulty node will re-establish a new LSP segment that bypasses the failure. An example is shown in Figure 4.11.

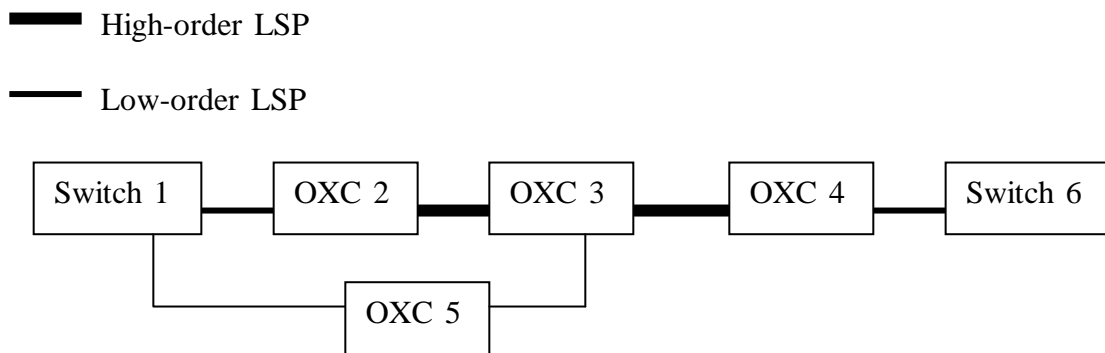


Figure 4.11: the situation where reroute Case 3 applies

In the example, a low-order LSP goes through (Switch 1, OXC 2, OXC 3, OXC 4, Switch 6). And the high-order LSP (FA-LSP) goes through (OXC 2, OXC 3, OXC 4). If OXC 2 is down, Switch 1 detects the neighbor failure, and it will initiate the reroute. It may trigger establishing another high-order LSP (FA-LSP) that bypasses OXC 2, e.g.,

FA-LSP (OXC5, OXC3, OXC4). And the low-order LSP is tunneled by the new FA-LSP.

The problem of multiple layer protection contention can also occur when using local restoration. For example, the link between two adjacent nodes is broken. If there is link layer protection there, e.g., that link is *Dedicated 1+1* protected link, it can provide faster recovery and the reroute should not be needed. So a coordination mechanism should be needed, e.g., the hold-off timer.

In the above cases, how does the reroute initiator node find the next hop to send out the signaling message so as to create the reroute path segment? The conventional RSVP [40] must consult the routing table. It expects that the changed topology has been shown in the routing table. But this does not happen right away after the fault in the network. So the conventional reroute to locally repair the link/node failure suffers packet loss. Let us see what is the problem.

When a link/node failure occurs in a network, routing protocols exchange the routing messages in the network to reflect a new topology. The routing information on different nodes may be temporarily inconsistent. And even a forwarding loop could be created. The situation causes packet loss. The longer the inconsistency lasts, the more packets are likely to be lost. The time consists of three periods: (1) the time the node needs to detect the failure, (2) the time a node needs to distribute the information across the network and (3) the time to reconstruct the routing table. Among these factors, period (2) is the major one (see [46]). To reduce the time it takes to detect link failure, we can use mechanisms in the link layer, e.g., in SONET, it is possible to detect link failure in less than 10 ms by SONET-specific mechanisms, such as Loss of Frame detection. But with regard to (2), the distribution nature of IP routing and the need for all the nodes to converge to consistent routing place limitations on how much (2) can be reduced. In practice, the time to converge within a single routing domain may be on the order of seconds. That means the packet loss may last on the order of seconds. Let us have an example (see Figure 4.12). Let us assume that the link between Router 1 and 2 is broken. Router 1 detects the failure. With the current routing table, Router 1 can find out that there is an alternative path (R1, R3, R4, R5, R2). Then, to create the reroute path segment, Router 1 sends out the RSVP-TE Path refresh message destined for Router 2 to Router 3. But due to the routing information distribution delay (2) mentioned above, Router 3 thinks R1 should be the next hop because it only takes 2 hops (R1 and R2). So Router 3 forwards the message back to Router 1. Thus the forwarding loop is created for a short period of time. The message is discarded. Packets would be lost for seconds and signaling fails until the next refresh time.

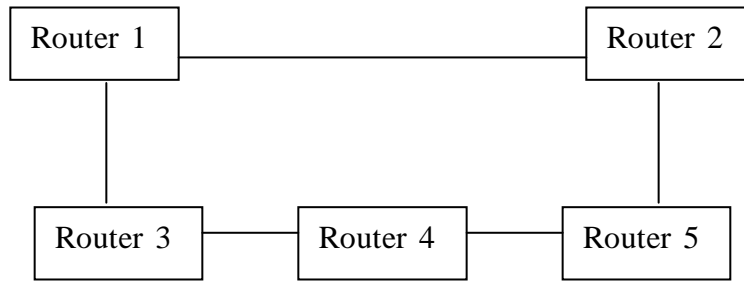


Figure 4.12: local reroute suffers packet loss for seconds

The conventional RSVP [40] suggests the signaling protocol wait a period of time, named *W*, before consulting the routing table to signal the bypass route. The recommended default value for *W* is 2 seconds. But this delay is not acceptable for many applications.

Yakov Rekhter and Bruce Davie in their book [47] suggest using an explicit-routed LSP as the reroute LSP segment to bypass the failure. Instead of using hop-by-hop, destination-based forwarding, the immediate upstream node of the faulty link/node constructs an explicit-routed LSP that bypasses the fault. The explicit-routed LSP merge with the original LSP at the node that is the immediate down stream node of the fault. Such an LSP uses the label stacking capability of MPLS to tunnel all the LSPs that used to going through the faulty link/node. And the rest of the original LSP does not need to be torn down or modified. Let us have an example to illustrate the idea.

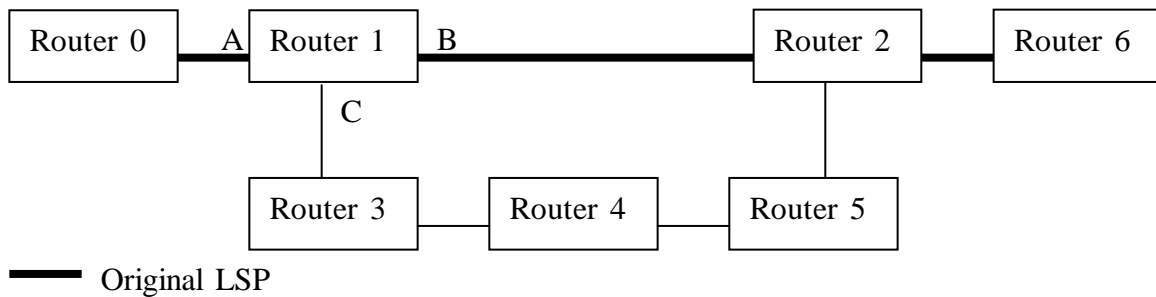


Figure 4.13: explicit-routed LSP bypasses the fault

In Figure 4.13, an LSP from Router 0 and Router 6 that is routed over Router 1 and 2. In the label forwarding table of Router 1 for that LSP, incoming label 10 and interface A corresponds to outgoing label 11 and outgoing interface B. It means, the packets from Router 0 with (incoming) label 10 through interface A will be replaced with 11 and forwarded to Router 2 through interface B. When Router 2 receives any packet with label 11 from Router 1, it will continue label forwarding, e.g., it forwards the packet to Router 6.

Let us assume that the link between Router 1 and Router 2 is broken. Router 1 detects the link failure, and it can construct an explicit-routed LSP right away, which is (Router 1, Router 3, Router 4, Router 5, Router 2), because its routing table tells it that there is such a route from Router 1 to Router 2. The topology change does not have an effect on constructing such an explicit-routed LSP. Now how to tunnel the original LSP? The Path message carries the ERO containing (Router 1, Router 3, Router 4, Router 5, Router 2), which specifies the explicit-routed LSP. The ERO drives Path message from Router 1, Router 3, Router 4, Router 5, and finally to Router 2. Router 2 responds with a Resv message, which allocates label 20 to Router 5. Similarly, Router 5 allocates label 21 to Router 4, Router 4 allocates label 22 to Router 3, and Router 3 allocates label 23 to Router 1. Router 1, receiving the label, re-programs the label forwarding table. First, it adds one more operation to the label forwarding process, which is to push label 23 on the packet that is from Router 0, and this operation is after replacing label 10 with 11 on the packet. Second, the outgoing interface is not B any more, but C. Router 2, as the ending node, may support penultimate hop popping.

Now, assuming that the packet with label 10 arrives at Router 1. Router 1 replaces label 10 with label 11 (as it did before the link failure), furthermore it pushes 23 on top of label 11. And it forwards the packet to interface C. Router 3 forwards the packet to Router 4 by replacing label 23 with 22. Similarly, Router 4 forwards the packet to Router 5 by replacing label 22 with 21. Router 5 forwards the packet to Router 2 by replacing label 21 with 20. When the packet arrives at Router 2, label 20 is striped off, and the label 11 becomes the top label. As before, Router 2 understands how to forward packets with label 11. The label allocation can be seen in Figure 4.14.

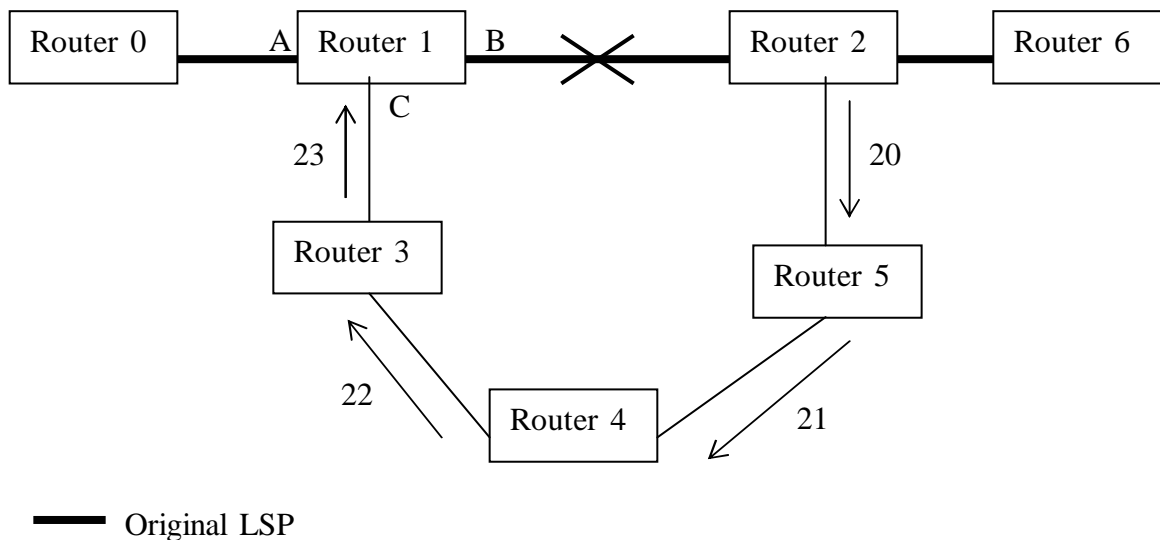


Figure 4.14: explicit-routed LSP bypasses the fault

The advantage of this solution is that the immediate upstream node of the fault does not need to wait for the routing information distribution or routing database synchronization. And it does not need to wait for reconstructing the routing table because the routing

database can still tell the reroute initiator node if there is another route to repair the path even after the link/node failure. Note that the reroute initiator node is the immediate upstream node of the failure. Thus, the waiting period of time (2) and (3) suffered by conventional IP routing can be eliminated. So this solution, which uses explicit-routed LSP to reroute, is faster.

From the above analysis, we can also see that this solution does not need to change the current signaling protocols, but it requires the nodes implement the intelligence to support this solution.

### **Summary of Local Rerouting**

Local restoration eliminates the need to propagate fault information across networks. But its application is limited.

As specified in Section 4.3.1 of this report, an explicitly routed LSP (ER-LSP) is pre-computed, which usually meets some traffic engineering goals. If a user's LSP is an ER-LSP, it is highly desired not to be rerouted. For example, the user's ER-LSP can route away from network congestion and bottlenecks. The local restoration mechanism works for the hop-by-hop routed LSP recovery very well, and it also works for the loosely specified portion of an ER-LSP, but not for a strictly routed ER-LSP. The local reroute mechanism is dynamic – it repairs the LSP by bypassing the failure after the failure occurs, and the new LSP travels some nodes/links that may not be desired. Such an LSP may not be optimal. Therefore, if we use local reroute mechanism for a user's ER-LSP, then after the local repair for a strict ER-LSP or the strictly specified portion of a loose ER-LSP, the initiator node of the LSP must be notified. And it should re-compute the LSP, and establish a new ER-LSP to meet the original requirements.

Using conventional local reroute takes a lot of time to wait for the routing information synchronization, and the local reroute using Yakov Rekhter and Bruce Davie's proposal (see [47]) provides a solution to solve the problem. But the signaling for creating the reroute path still takes time.

Because of the network topology, local repair may not succeed.

### **5.2.2 Global Restoration**

With global (end-to-end) path restoration, the backup path is not established until the failure on the path occurs. After the failure is detected, the initiator node of the faulty LSP is notified of the failure. And it establishes the alternate path destined for the destination node and switches the traffic to the new path.

When a failure occurs, the fault detection triggers the fault localization mechanism. After the location of the fault is found, the node that is closest to the failure distributes the fault information by the routing protocol. In the meantime, it notifies the node that initiates the LSP establishment.

The faulty LSP should be torn down and the resource allocated for the faulty LSP should be freed. The information is also distributed by the routing protocol.

The LSP initiator node should wait for the routing information synchronization. After that, it re-establishes another LSP that bypasses the failure and the traffic is switched over onto the new LSP.

### Summary of Global Restoration

Compared to end-to-end path protection, the end-to-end path restoration is slow because the fault notification and the routing information synchronization would take seconds. So it does not work for real-time applications such as voice. It is resource efficient, because the alternative LSP is established on demand and the resource is allocated on demand.

In order to eliminate the time for routing information synchronization, Yakov Rekhter and Bruce Davie’s proposal (see [47]) can be also used for end-to-end path restoration.

## 6. Case Studies

### 6.1 Case Study 1: The end-to-end LSP Protection

The network topology is shown in Figure 5.1. The switches in the client networks are SONET switches and the OXCs in the optical core network operate on a single lambda level. Let us assume that the edge OXCs have the capability to convert electrical signals to optical signals. They have interfaces that can provide SONET signals and interfaces that can provide WDM capability. There are two OC192 links that connect edge nodes, e.g., SONET switch S3 and OXC O1. The optical fiber between two OXCs can contain 16 lambdas, each of which can provide capacity equivalent to one OC192 capacity.

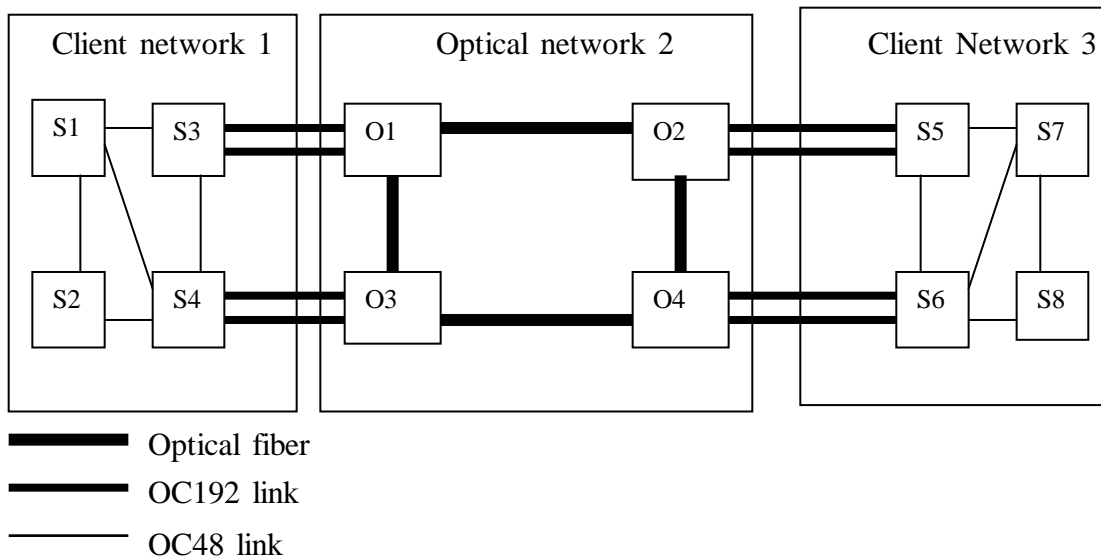


Figure 5.1: the network to show end-to-end 1:1 LSP protection

A client LSP is going to be established between Switch 1 of client network 1 and Switch 7 of client network 3, which requires 1:1 LSP protection. Switch 1 is the client LSP

initiator. This client LSP requires 622Mb/s (OC-12) bandwidth, and it is required to use links whose administration color is “red”. Note that a link is usually colored to indicate which administration group it belongs to. Here we assume that the client wants the links that belong to the administration group “red”.

To support traffic engineering, the primary path is an ER-LSP and it is pre-computed. Because the primary and the backup path are disjoint, the backup path should be also pre-computed. The database (TE-LSDB) stores the link TE information of the network. Let us assume that it has the information in Table 5.1 (Table 5.1 is on the coming page). For simplicity, the data encoding type contained by the Interface Switching Capability Descriptor (ISCD) is ignored. We only consider the interface switching type and the maximum reservable bandwidth of the ISCD. And we also assume that the TE information is the same for both directions of a link.

Because of the administration constraint, we only consider “red” links. So link (S1, S4) is excluded. Link (S6, S7) only has 500 Mb/s available, which is less than the required bandwidth. So it is also not considered. If we use a link whose link protection type is not “unprotected”, e.g., “dedicated 1+1”, then we must configure the coordination mechanism at each node of the path so as to avoid multiple-layer protection contention. If a node has intelligence to configure itself (e.g., the auto-configuration mechanism), then manual configuration is not needed. For example, we set a policy in each node – if the link protection type is not “unprotected”, then the lower layer protection has higher priority and the hold-off timer is one second. When the signaling message arrives at the node, the node configures itself based on link protection type. Another choice is to only use links whose link protection type is “unprotected”, and disable the link layer protection (e.g., set the hold-off timer to zero). Let us take this choice. Therefore, the topology we will consider becomes as in Figure 5.2. We calculate the metric (cost) of the link by  $(10^8 / \text{available bandwidth})$  and we have the cost of the link, which is also shown in Figure 5.2.

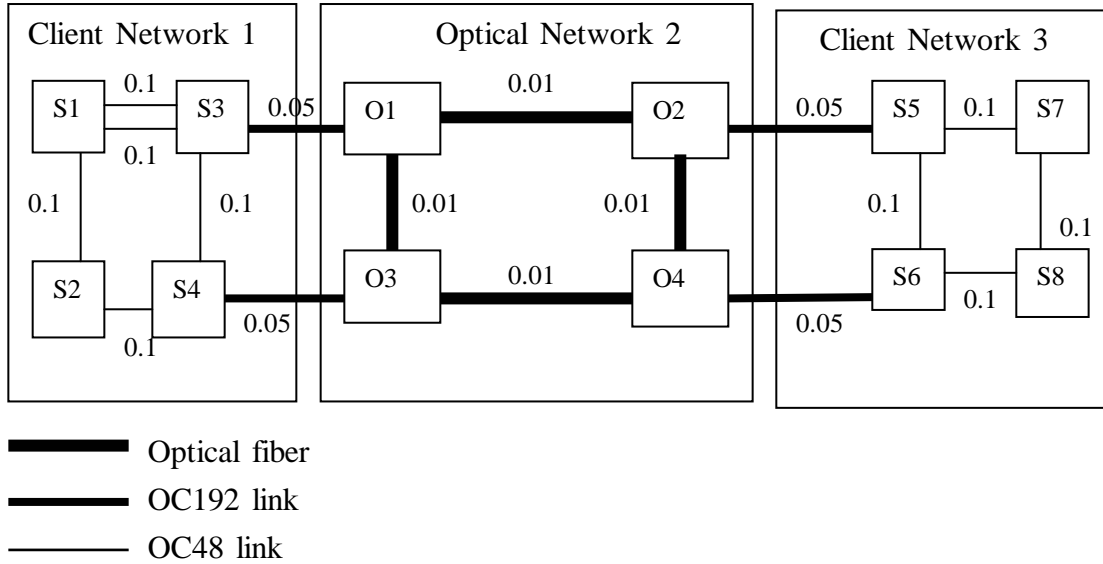


Figure 5.2: the topology that does not violate the constraints

Links (S1-1, S3-1) and (S1-2, S3-2) are equivalent links. By some policy (e.g., random), the first one is chosen. Then we can use the SPF algorithm to calculate the “shortest” path, which is (S1-1, S3-1, O1, O2, S5, S7). This is the primary LSP.

**Legend in Table 5.1**

- SRLG: Shared Risk Link Group
- ISCD: Interface Switching Capability Descriptor
- MRB: Maximum Reservable Bandwidth



Local Address	Remote Address	SRLG	ISCD	Unreserved Bandwidth	Link Protection Type	Admin. Color
S1-1	S3-1	11	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S1-2	S3-2	11	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S1	S2	12	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S1	S4	13	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	green
S2	S4	14	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S3	S4	15	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S5	S7	31	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S5	S6	32	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S6	S7	33	TDM, MRB =2.5Gb/s	500Mb/s	unprotected	green
S6	S8	34	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S7	S8	35	TDM, MRB =2.5Gb/s	1 Gb/s	unprotected	red
S5-1	O2-1	231	TDM, MRB =10Gb/s	2 Gb/s	unprotected	red
S5-2	O2-2	231	TDM, MRB =10Gb/s	2 Gb/s	Dedicated 1+1	red
S6-1	O4-1	232	TDM, MRB =10Gb/s	2 Gb/s	unprotected	red
S6-2	O4-2	232	TDM, MRB =10Gb/s	2 Gb/s	Dedicated 1+1	red
S3-1	O1-1	121	TDM, MRB =10Gb/s	2 Gb/s	unprotected	red
S3-2	O1-2	121	TDM, MRB =10Gb/s	2 Gb/s	Dedicated 1+1	red
S4-1	O3-1	122	TDM, MRB =10Gb/s	2 Gb/s	unprotected	red
S4-2	O3-2	122	TDM, MRB =10Gb/s	2 Gb/s	Dedicated 1+1	red
O1	O2	21	LSC, MRB =160 Gb/s	10 Gb/s	unprotected	red
O1	O3	22	LSC, MRB =160 Gb/s	10 Gb/s	unprotected	red
O2	O4	23	LSC, MRB =160 Gb/s	10 Gb/s	unprotected	red
O3	O4	24	LSC, MRB =160 Gb/s	10 Gb/s	unprotected	red

Table 5.1: the TE link information for path calculation

Because link (S1-2, S3-2) has the same SRLG as the link (S1-1, S3-1), and the latter has been chosen for the primary LSP, it should not be considered when calculating the backup LSP. The primary and the backup LSPs should be disjoint, so the topology we can consider for the backup LSP becomes:

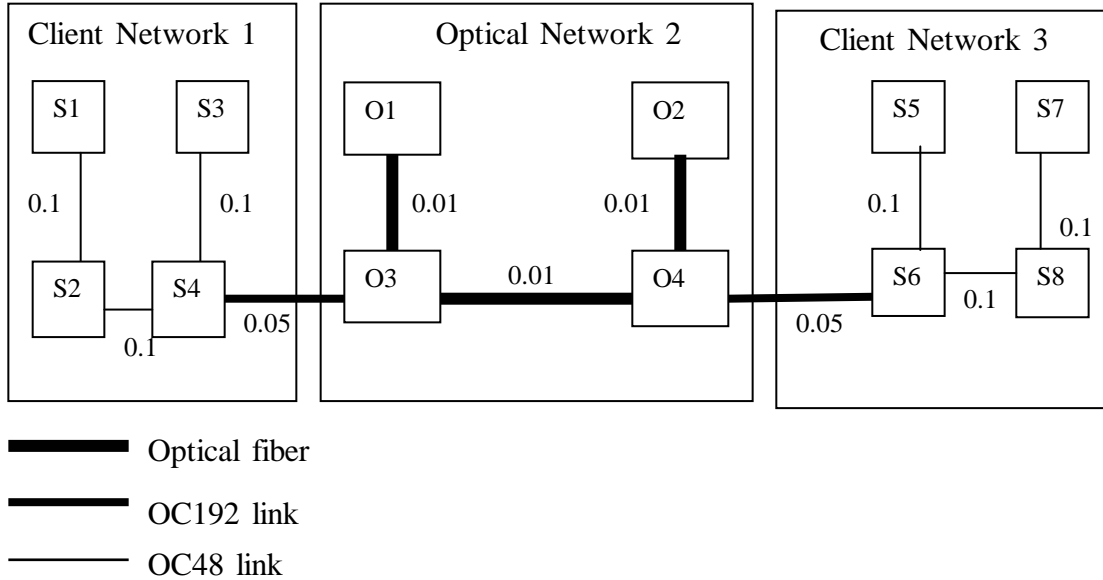


Figure 5.3: the topology for the backup LSP

Using the SPF algorithm, we have the “shortest” path (S1, S2, S4, O3, O4, S6, S8, S7) for backup LSP.

Now the Switch 1 can signal both LSPs, e.g., using RSVP-TE. Explicit-routed LSP (ER-LSP) signaling is used. The signaling protocol carrying the ERO establishes the LSP starting from Switch 1. When signaling arrives at SONET Switch 3, Switch 3 finds out that it is at the boundary for a hierarchical LSP by the Interface Switching Capability Descriptor. Let us assume that there is no existing FA-LSP that meets the requirement of the LSP being set up. So Switch 3 establishes a new FA-LSP starting from Switch 3 and terminating on Switch 5. Switch 3 initiates the new FA-LSP. When the signaling arrives at OXC1, OXC 1 finds out it also needs a new FA-LSP between OXC 1 and 2. Let us call this FA-LSP  $F1w$ , which has Link Protection Type “unprotected”. After that, FA-LSP between Switch 3 and 5 is tunneled through  $F1w$ . We call this FA-LSP (between Switch3 and Switch5)  $F2w$ . It also has Link Protection Type “unprotected”. This FA-LSP tunnels the client LSP. Eventually, the client LSP between Switch 1 and Switch 7 is established. It is the primary path we want, which we call  $Pw$ . A hierarchical LSP establishment is illustrated in Section 2 of this report.

Similarly, for the backup LSP, the FA-LSP between OXC 3 and 4 is called  $F1b$ ; the FA-LSP between SONET Switch 4 and 6 is called  $F2b$ . Both of them have Link Protection Type “unprotected”. And the backup LSP is tunneled through these FA-LSPs, and let us call it  $Pb$  (see Figure 5.4). LSP  $Pw$  and  $Pb$  construct the 1:1 LSP protection as desired.

The FA-LSP  $F1w/b$  and  $F2w/b$ , which have Link Protection Type “unprotected”, will be advertised by the routing protocol. And their unreserved bandwidth is the difference between the maximum reservable bandwidth and the share used for LSP  $Pw$  (or  $Pb$ ). For example, the FA-LSP  $F2w$  advertises that it has bandwidth 9.178 Gb/s available, and the FA-LSP  $F1w$  advertises that it has 15 lambdas available, each of which has OC192 bandwidth.

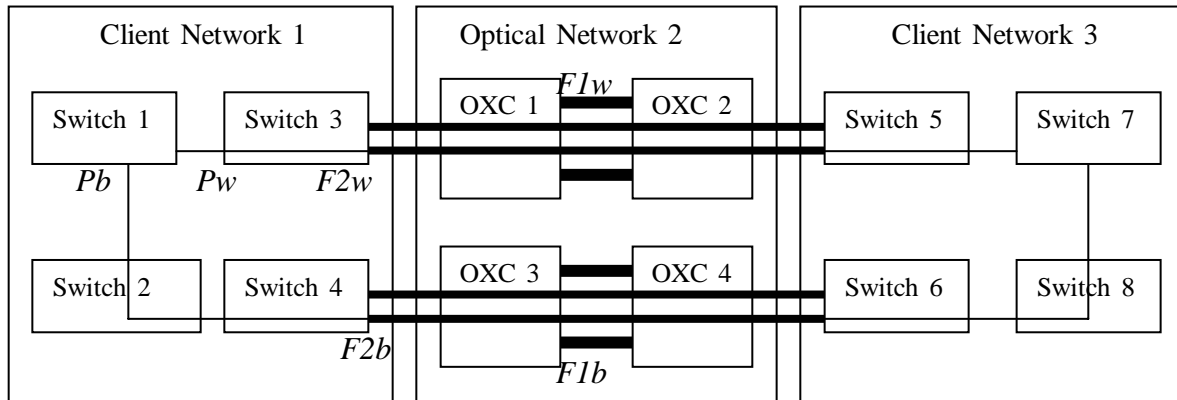


Figure 5.4: the 1:1 LSP protection

During the signaling, the resources are reserved. When the signaling takes place, the RSVP-TE Path message carries a flag that tells the nodes the LSP being signaled is the primary LSP or the backup one. Because the user requires 1:1 LSP protection, the user’s traffic is not transported over the backup LSP until a failure occurs. The resource of the backup LSP may be used by other LSPs that have lower priorities.

When the RSVP-TE Path message is sent out, it carries the Notify Request object. It has the “targeted” node IP address, which is Switch 1 in this case. Every node along the path records this IP address. This is the end-to-end LSP protection. It is not necessary for the node that is responsible for triggering the traffic switch to know exactly where the failure occurs on the path. So it is not necessary to localize the failure. All nodes that detect the failure report the failure to the LSP initiator node. They send out a RSVP-TE Notify message destined for the targeted node – Switch 1. The LSP initiator node can trigger the traffic switch as soon as it receives the first notification, e.g., even one RSVP-TE Notify message.

Such an LSP protection can protect any failure on the LSP. But it takes a long time for the fault notification to travel the networks to reach the LSP protection initiator. For many real-time applications, e.g., voice over IP, it is highly desirable to be able to recover in 10s of milliseconds [48]. Fault notification may not work so fast. Therefore protection needs to improve for real-time applications. If what the user requires is the end-to-end restoration, the protection LSP is not pre-established. The primary LSP initiator does not start signaling the protection LSP until the failure occurs and the initiator is notified. So

end-to-end restoration is even slower and obviously it does not meet the requirement of real-time applications.

We will see how another end-to-end protection scheme can improve the recovery time in the next section.

## 6.2 Case Study 2: The Domain-specific Protection

A recent proposal [49] describes a GMPLS LSP protection scheme that is based on different network domains. It is called *subnetwork protection*.

The network across which a hierarchical LSP travels is partitioned into subnetworks. The nodes constructing the subnetwork have the same multiplexing capacity. Within each subnetwork, there is a pre-established backup LSP to protect the primary LSP. And the resource may also be pre-allocated. Because all nodes in a subnetwork have the same multiplexing capacity, the primary and the backup LSP are at the same level in the LSP hierarchy. The protection mechanism in each subnetwork can be M:N or 1+1. If there is a failure, for M:N protection mechanism, the traffic switchover occurs from the primary LSP to the backup LSP; for 1+1 protection mechanism, the LSP terminator node selects the traffic from the backup LSP. The protection is only performed within the subnetwork where the failure occurs. There is no need to do anything in other subnetworks across which the hierarchical LSP travels. The logical view of this idea is shown by the 1:1 protection mechanism in Figure 5.5. There is a protection LSP in the subnetwork for the primary LSP segment that goes over that subnetwork.

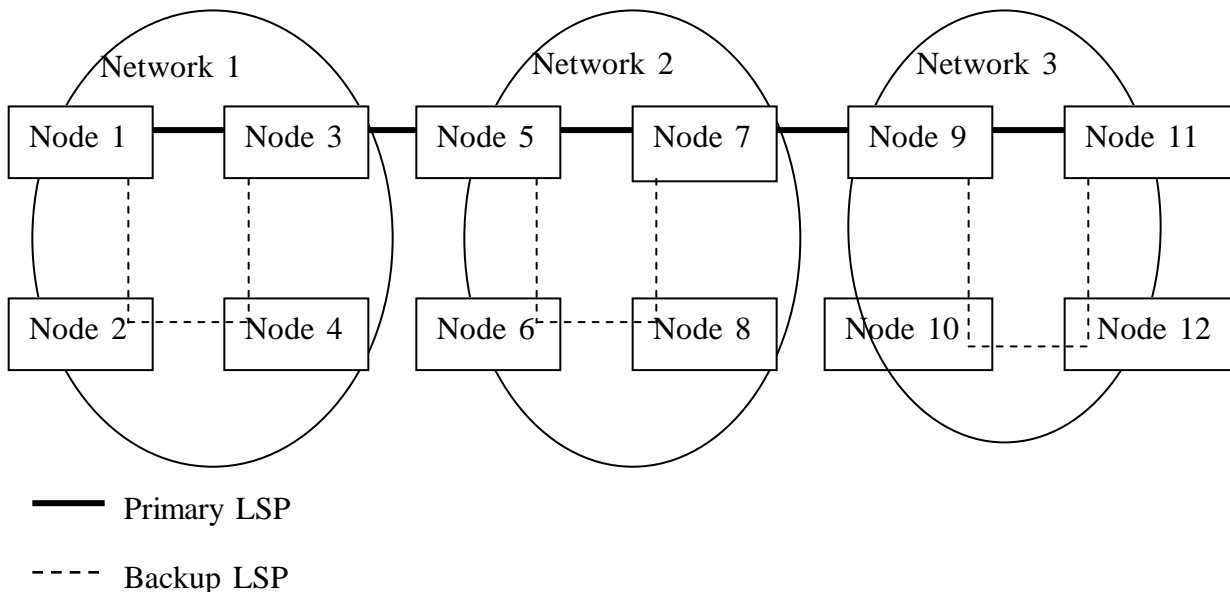


Figure 5.5: The logical view of the subnetwork protection

In Figure 5.5, the primary LSP is (Node 1, Node 3, Node 5, Node 7, Node 9, Node 11). If the link between Node 5 and Node 7 is broken, the protection LSP (Node 5, Node 6,

Node 8, Node 7) takes over the traffic, and there is no action in other networks. Traffic goes from Node 1 to Node 11 by (Node1, Node3, Node 5, Node 6, Node 8, Node 7, Node 9, Node 11).

In this subnetwork protection mechanism, the segments of the primary LSP are protected by the protection LSPs in different subnetworks. Compared to end-to-end LSP protection introduced in the previous section, this protection mechanism requires shorter time for fault notification as the fault notification only travels to the nodes within a subnetwork. Compared to local reroute, it is simpler. But, such a protection mechanism does not protect the nodes/links that are at the border of the subnetworks. The links at the borders can be protected by the link layer mechanism. However, the border nodes do not have protection. For example, there is no protection if Node 5 goes down in Figure 5.4. Fortunately, in practice, usually the nodes at the border of the network are very powerful and reliable.

Let us see how to implement such a protection scheme for the case we mentioned in the previous section. The client LSP has the same requirements as that in the previous section. Here let us re-use the network shown in Figure 5.1. Because the link (S6, S7) does not have enough available bandwidth and link (S1, S4) violates the constraint, we have the topology as in Figure 5.6 to consider.

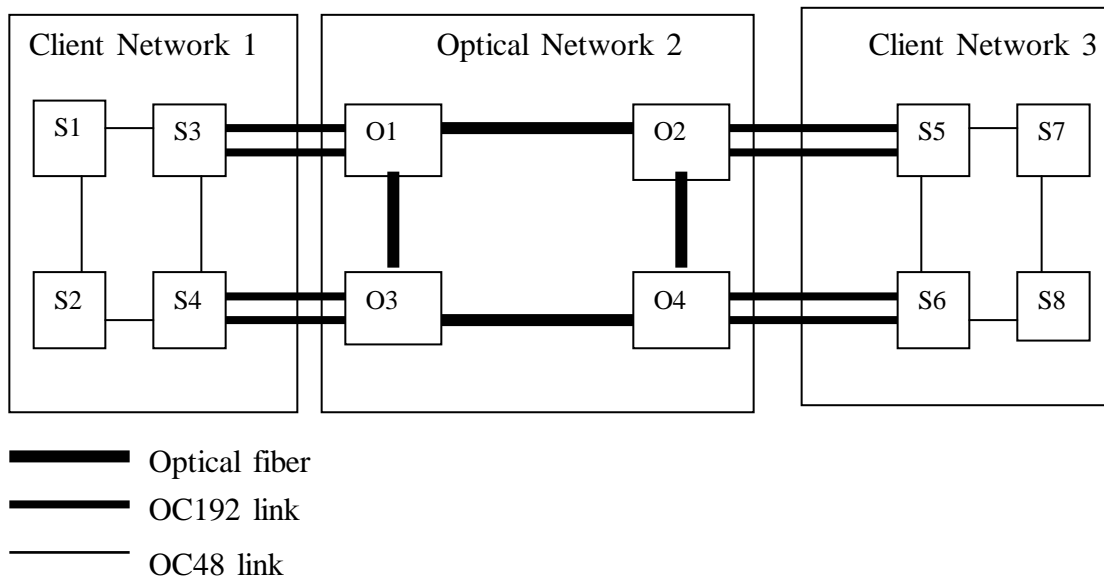


Figure 5.6: the network to show the subnetwork LSP protection

After the implementation, the LSP protection should be as follows. Switch 1 initiates the LSP, and this client LSP is tunneled by the high-order LSP from Switch 3 to Switch 5, which in turn is tunneled by the higher-order LSP from OXC 1 to OXC 2. Finally the client's LSP terminates at Switch 7. Because the 1:1 LSP protection is required, we choose such a method - all the links within the networks have link protection type "unprotected", but the link between Switch 3 and OXC 1 has link protection type

“Dedicated 1:1”, e.g., the SONET APS link layer protection. So is the link between OXC 2 and Switch 5. The primary LSP is (link (S1, S3), link (S3-2, O1-2), link (O1, O2), link (O2-2, S5-2), link (S5, S7)). Within Client Network 1, the LSP segment from Switch 1 and 3 is protected by LSP (Switch 1, Switch 2, Switch 4, Switch 3). Within the optical network, the LSP segment from OXC 1 to OXC 2 is protected by LSP (OXC 1, OXC 3, OXC 4, OXC 2). And within network 3, the LSP segment from Switch 5 to Switch 7 is protected by LSP (Switch 5, Switch 6, Switch 8, Switch 7). The resource has been allocated on these protection LSPs, but the protection LSPs do not transport traffic. Thus, the entire user LSP has 1:1 LSP protection except the edge nodes, like, OXC 1, OXC 2, Switch 3 and Switch 5, and except the initiator and terminator nodes (see the following figure). The protection mechanism can protect any failures between the edge nodes within each subnetwork.

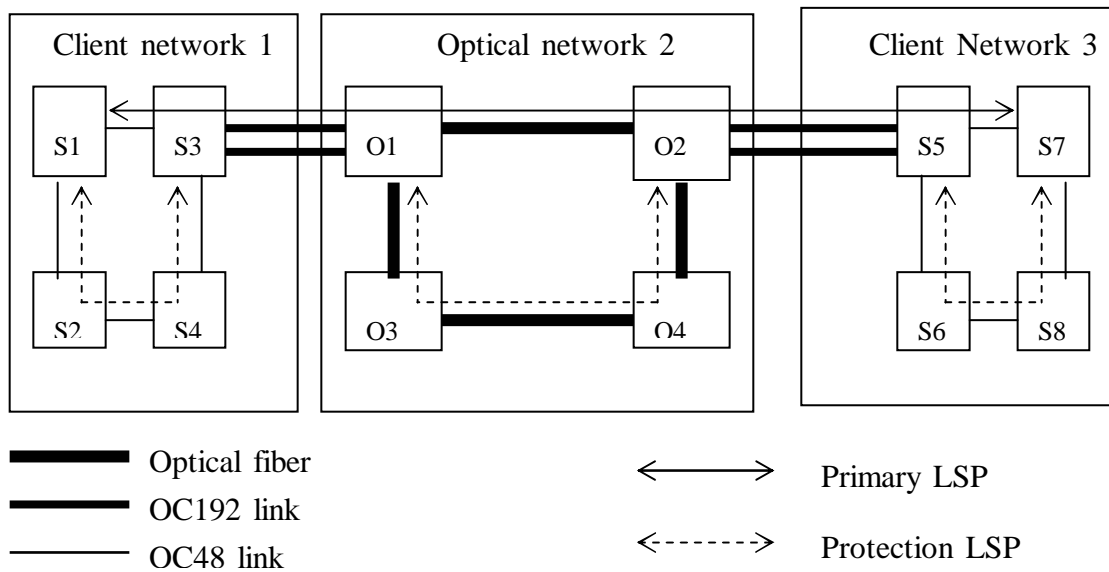


Figure 5.7: the network to show the subnetwork LSP protection

How to signal such a protection scheme? As this report is being written, there is no automatic mechanism proposed in IETF yet. Let us discuss what we need to do.

(1) The primary LSP and protection LSP should be disjoint within each subnetwork. It means the protection LSPs must be pre-computed, so they are explicit-routed LSPs

(2) Different protection mechanisms should be allowed within the subnetwork, e.g., 1+1 or 1:1. The LSP initiator node can be configured to create one of these protection mechanisms, but how to tell the ingress node (a node at which the working LSP enters a subnetwork) about the desired protection mechanism so that the ingress node signals the protection LSP? For example, in Figure 5.6 (on the previous page), how does the signaling protocol tell OXC 1 or Switch 5 to establish the protection LSP? And which protection mechanism is wanted, e.g., 1+1 or 1:1? The current signaling protocols do not provide any support yet, but it is possible to add some extensions to support this

*subnetwork protection* scheme, e.g., a new object in RSVP-TE. This new object is only processed by the nodes of the primary LSP that are at the border of different subnetworks. For example, Switch 1, Switch 3, OXC 1, OXC 2, Switch 5, Switch 7 in Figure 5.6 (see the previous page).

(3) The protection LSP should be pre-established so as to provide fast recovery. Resources may be pre-allocated as well. For M:N protection, lower priority traffic should be allowed to use the resource if the protection LSP is not protecting.

(4) Coordination mechanisms should be used to avoid the multi-layer protection contention if there is any. For example, “unprotected” link protection type may be used to signal both of the primary and backup LSPs.

(5) There is a problem concerning the incoming interface. Within each subnetwork, the primary LSP segment and the backup LSP merge at the edge node. The incoming interface may be regarded as a “label” and involved in label switching, e.g., in a network constructed by nodes that is fiber-switch capable, the incoming port may determine the outgoing port. Another example is an MPLS router that is packet-switch capable uses interface-based label space. The problem is illustrated in the following figure.

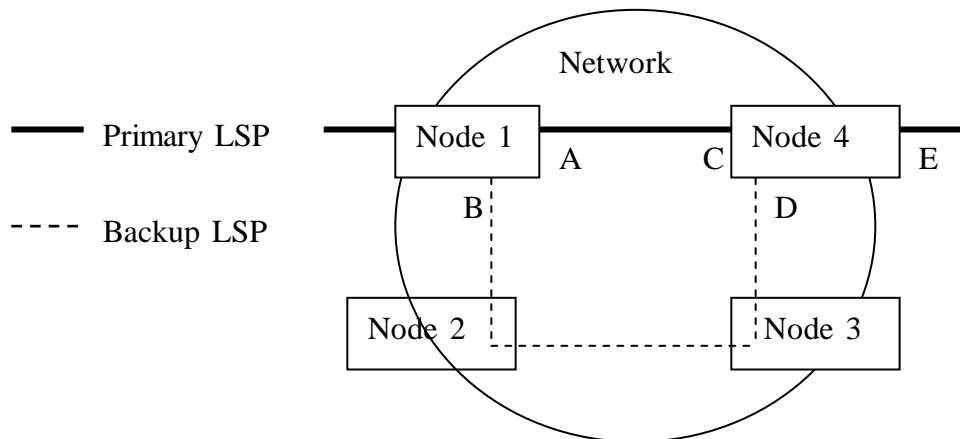


Figure 5.8: the incoming interface problem in the subnetwork protection

Node 1 switches the traffic from interface A to interface B if a failure between itself and Node 4 occurs. Then the traffic arrives at Node 4 through interface D, instead of C. If the label is unique node-widely (per-node label space), then there is no problem for Node 4 to work as usual, and the discussion can be stopped here. But in many situations, this is not the case. In order to reuse the label, usually per-interface label space is used. For example, a fiber can transport multiple wavelengths (lambdas), and another fiber on a different port can transport all the same wavelengths (lambdas). Let us assume that Node 4 has such an entry (see Figure 5.9) in its label forwarding table in the example shown in Figure 5.8:

Incoming information	Outgoing information
Incoming label	Outgoing label
Incoming interface C	Outgoing interface E
	...

Figure 5.9: the label forwarding entry in the example

Now the incoming interface has changed for Node 4, and how to tell it to accept the traffic from another interface and continue the label forwarding? One solution is to signal Node 4 to change the incoming interface C to D in its label forwarding entry after the failure is detected. It takes time and this protection would lose much of its value. Another solution is to tell Node 4 about it when the protection LSP is being established. In order to support this solution, a selector may be implemented in Node 4 that can select traffic from one of the multiple ports. Node 4 monitors the traffic from interface C and D, and it selects the healthier traffic from one of the two. The incoming interface may be programmed in the label forwarding entry before label switching occurs for optimization (see Figure 5.10) if the protection type allows. Or Node 4 can change the incoming interface in its label forwarding entry just before it is going to select the traffic from another interface.

Incoming information	Outgoing information
Incoming label	Outgoing label
Incoming interface C	Outgoing interface E
Incoming interface D	...

Figure 5.10: the interfaces for primary and backup LSPs are pre-programmed

(6) How to set up the multi-layer protection scheme like the link layer protection between nodes S3 and O1, O2 and S5? It is done usually by configuration. So is the set-up of coordination mechanism to avoid the multi-layer protection contention.

The other way to establish the entire *subnetwork protection* is by configuration. For example, on the network manager, the network administrator can configure such an LSP protection scheme. At the beginning, the network administrator requires the path computation component in the primary LSP initiator node to calculate the primary LSP. Then the LSP protection type and the primary LSP information (e.g., the nodes traveled



by the primary LSP) are sent to the ingress node of the primary LSP segment within each subnetwork, for example, node O1 in the optical network in Figure 5.6. At each ingress node, the protection LSP is calculated to protect the LSP segment that travels within that subnetwork. Note that the protection LSP must be disjoint with the primary LSP segment and the protection type should be honored. After that, the link layer protection (if needed) and the coordination for avoiding multi-layer protection contention can be done by configuration. How to solve the incoming interface problem? The egress node may provide an interface to network management for query and configuration. Such an interface allows the network administrator to manually query and configure the label forwarding table. We can see that using configuration to create such a protection scheme is tedious and error-prone.

### Summary of the Subnetwork Protection Scheme

If the link between subnetworks fails, then the link layer protection is triggered. And it is expected that the link layer protection takes a short time to recover, e.g., the SONET APS just takes less than 50 ms to recover. If there is a failure (not the edge nodes) in a subnetwork, fault notification just needs to notify the head node of the LSP segment within that network. So the notification message travels only within that subnetwork. Compared to end-to-end LSP protection, it takes less time. The paper [50] proves that, in theory, it is possible to guarantee the 50 ms recovery time in large mesh networks by properly partitioning the network and applying subnetwork protection.

This subnetwork protection scheme also has another advance – it can protect a number of LSPs (see Figure 5.11). If a failure between Node 1 and Node 4 occurs, the protection LSP, which has the same level as the primary LSP segment within the subnetwork, is activated to protect the primary LSP. The tunneled low-order LSPs, e.g., LSP 1, 2 and 3 in the example, are not affected, and they are not even aware of the failure.

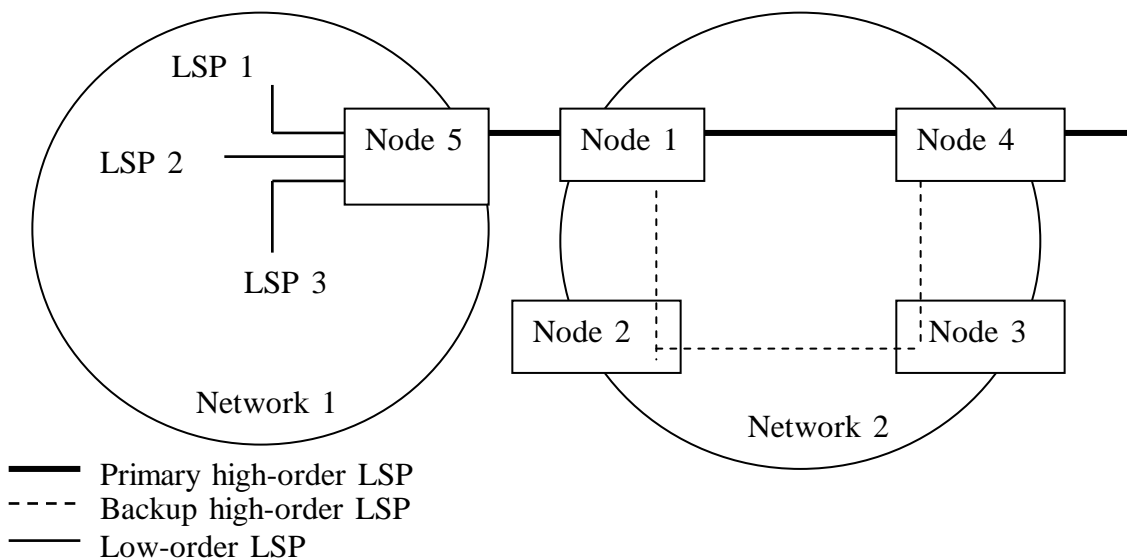


Figure 5.11: the subnetwork protection idea protects multiple low-order LSPs

This subnetwork protection scheme is resource-efficient. For example, the dedicated 1:1 end-to-end LSP protection mechanism doubles the resource. But in the subnetwork protection scheme, the resource for 1:1 LSP protection is shared - the protection LSP can be shared by multiple low-order LSPs.

Compared to local/global restoration, the protection LSP in the *subnetwork protection* is pre-established. So it provides faster recovery. But as other protection mechanisms, it requires more resource than restoration.

The signaling issues to solve the incoming interface problem in this subnetwork protection scheme needs further study.

### 6.3 Case Study 3: Link-layer Protection and Local Reroute

In the mesh network shown in Figure 5.12, photonic switches construct the core network. At the edge, devices O1 and O2 are optical switches. The optical switch has interfaces that provide WDM capabilities for photonic switches, and interfaces that provide SONET section level signals. SONET switches are connected to O1 and O2. They provide OC-192 capacity interface. Between O1 and P1, it is the WDM multiplexing of 16 OC-192 signals which remain intact through to O2. All lines have dedicated 1+1 link protection (the dedicated protection link is not shown in the figure). The links between SONET switches are OC48 links, like the link between S1 and S3, the link between S5 and S7. The optical switches O1, O2, O3 and O4 are IP-over-WDM nodes. So are the photonic switches.

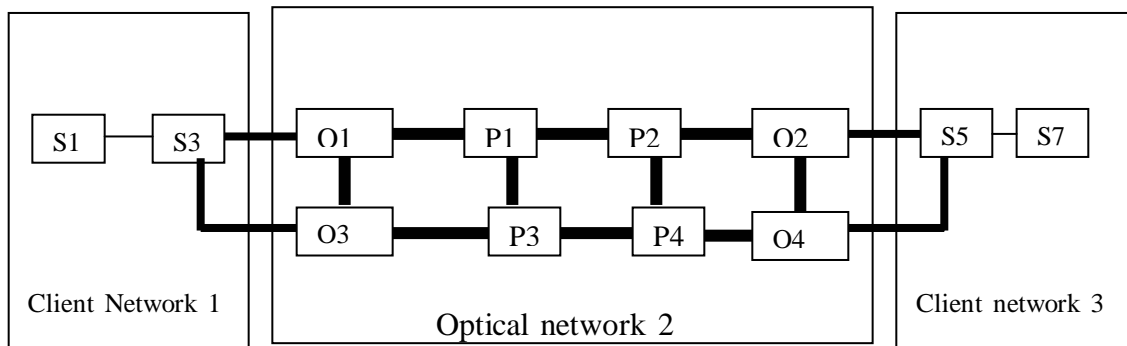


Figure 5.12: an LSP requiring 1+1 protection is built in the mesh network

A client LSP is going to be established between Switch 1 of client network 1 and Switch 7 of client network 3. It requires fault recovery in the optical network. The LSP will be used to transport real-time applications and the recovery should be done quickly if there is a failure, e.g., in 10s of milliseconds.

Link layer protection is one of the solutions for fast failure recovery. Let us study it here to see if 1+1 link layer protection can work in this case. If we build an LSP whose links all have “Dedicated 1+1” link layer protection type, the whole LSP has link protection. But what happens if a node goes down? Let us see an example in Figure 5.13. All the

nodes are IP-over-WDM nodes. If node N3 goes down, how to recover the failure even if all the links have 1+1 link protection? So just link layer protection cannot work. Other recovery mechanisms are needed to complement the link protection.

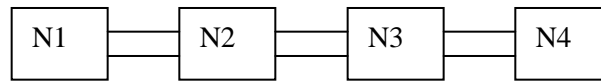


Figure 5.13: all links have 1+1 link protection between nodes

Because an LSP which has 1+1 link protection has doubled the resource for the traffic, further expensive recovery mechanisms are not desired any more. One of the solutions is to use local reroute. Let us consider if this recovery mechanism can work.

When establishing the primary LSP, the RSVP-TE Path message carries the RSVP-TE protection object, which signals “Dedicated 1+1”. To avoid multiple layer protection contention, the coordination mechanism must be set during the signaling. Let us use hold-off timer. Note that, in order to meet the recovery time requirement, the hold-off time set-up must consider the time needed for MPLS-based recovery in case the link layer protection fails. We use LSP local reroute as the MPLS-based recovery in this case. And the link layer protection has higher priority than the MPLS-based recovery. Let us assume that the primary LSP is (S1, S3, O1, P1, P2, O2, S5, S7). The primary LSP contains: FA-LSP1, which is from O1 to O2; and FA-LSP2, which is from S3 to S5. When establishing FA-LSP1, P2 knows that it is the penultimate node of this FA-LSP, e.g., routing tells P2 that it is directly connected to O2. Let us assume we have such an administration policy that the penultimate node of the FA-LSP must notify the initiator node of the FA-LSP. The “target” address for the Notify message can be configured. In this case, P2 can send out the RSVP-TE Notify message targeted to O1. O2 is the penultimate node of FA-LSP2, and similarly it knows it will send a Notify message to S3 if S5 fails.



Figure 5.12: the primary LSP

At first, we consider Case 1 (see the section about local restoration for what the different cases are), for example, P2 goes down. P1 detects its neighbor’s failure, e.g., by the Hello protocol (the Hello is exchanged between the neighbors every 5 ms). The link layer protection is triggered. Unfortunately, after the hold-off time, P1 finds out the failure is still there. So the hold-off timer triggers the LSP local restoration. The routing information database (LSDB) in P1 still shows that there is a route (P1, P3, P4, O2) to O2. Without waiting for routing information synchronization, P1 constructs an ER-LSP to reach O2, whose RSVP-TE ERO object contains P1, P3, P4 and O2. Because all of interfaces connecting these nodes have the same interface switching type – Lambda Switch Capable, there is no higher-order LSP needed. The reroute ER-LSP (P1, P3, P4, O2) has the same level as FA-LSP1 (O1, P1, P2, O2). When the signaling RSVP-TE

Path message driven by the ERO object arrives at O2, based on the RSVP-TE Session object and Sender template object, O2 understands the LSP has to be modified. So O2 modifies its label forwarding table and responds with a RSVP-TE Resv message. The message arrives at P1. And P1 understands that the reroute succeeds. So it also modifies its label forwarding table and switches over the traffic onto the reroute ER-LSP. If node P1 goes down, the reroute process is similar as both P1 and P2 are transit nodes of FA-LSP1.

If O2 fails, then reroute Case 2 occurs. P2 detects its neighbor's failure. As P2 is configured to notify the FA-LSP1 initiator O1, it sends out the RSVP-TE Notify message destined to O1. O1 is notified, and it tells all the tunneled low-order LSPs to reroute as it is the border of the hierarchical LSP. For example, it tells node S3 of FA-LSP2 to reroute. S3 consults its current routing database and builds the ERO object to signal a reroute ER-LSP. It understands it must cross the optical network to reach SONET switch S5. So the ERO object (S3, O3, P3, P4, O4, S5) is built and part of this ER-LSP (O3, P3, P4, O4) is a higher order LSP compared to FA-LSP2. The ERO drives the signaling. When it arrives at node O3, the higher-order FA-LSP is triggered to set up – let us call it FA-LSP1'. After that the reroute ER-LSP reaches S5. And the FA-LSP2 is tunneled by this FA-LSP1'. The reroute bypasses the faulty O2.

If O1 fails, the reroute Case 3 occurs. S3 detects its neighbor's failure and S3 triggers the reroute. S3 consults its current routing database and builds the ERO object to signal a reroute ER-LSP. The process is like what the S3 does in reroute Case 2 (see the preceding paragraph).

When we use reroute as the recovery method, we need to carefully consider the network topology. Due to the network topology, reroute may not work. For example, in the case we just described, if the user wants the fault recovery from end to end, reroute cannot work if node S5 goes down.

## 7. Conclusion

We have talked about the objectives for the LSP protection/restoration in Section 3. We note that the objective to *be cost-effective* may involve non-technical factors, but we do not discuss them here in this report. We compare the LSP protection/restoration mechanisms in GMPLS networks in the following table.

LSP recovery mechanisms	Resource requirements	Speed of recovery	Complexity	Application scope
Conventional local restoration	No resource is pre-allocated, the repaired LSP requires same resource	Very slow as it waits for the routing synchronization	No change to the current signaling protocols	Limited as the user's strict ER-LSP is not desired to be rerouted.
Local restoration with ER-LSP [47]	No resource is pre-allocated, the repaired LSP requires same resource	Fast. It does not wait for the routing synchronization to signal the reroute path. The path computation takes little time.	No change to the current signaling protocol, but it requires extra intelligence	Limited as the user's strict ER-LSP is not desired to be rerouted.
End-to-end restoration	No resource is pre-allocated, the repaired LSP requires same resource	Very slow. Fault localization is performed, fault notification takes time to travel across networks, and the reroute LSP is not set up until the failure occurs.	No change to the current signaling protocol	Can be used in any situations and the recovery meets traffic engineering goals
Local protection	Double resource is pre-allocated	Very fast as it is done at the link/physical layer	Additional configuration is needed to set up	It cannot easily provide node protection.
End-to-end protection	Additional resource is pre-allocated, dedicated 1+1 LSP protection requires double resource	1+1 LSP protection does not need fault notification but M:N LSP protection does.	Additional configuration may be needed to set up the protection on the end nodes of the LSP	It can be used in any situations

Table 6.1: comparison of recovery mechanisms

All protection/restoration mechanisms sacrifice resource to achieve fast recovery. Because additional resource is pre-allocated in the protection mechanism, the protection mechanism is expected to provide faster recovery than restoration. So objectives for LSP recovery

(1) *to optimize the use of resources* and (2) *to provide fast recovery and minimize the disruption to data traffic of any failure* are conflicting. Many protection/restoration mechanisms require signaling at the time of failure. The more signaling is required, the more time the mechanism takes to recover, and the less likely the recovery is timely.

We can achieve fastest recovery if we pay double resource, e.g., using the link/physical layer protection. The 1+1 LSP protection requires double resource, which is the most expensive LSP protection, and it can provide fastest LSP recovery. Any other protection mechanisms that share backup resource require fault notification. For example, the M:N, 1:N or 1:1 end-to-end LSP protection requires that fault notification travels across a number of nodes, which may cost time. The subnetwork protection mechanism tries to shorten the fault notification time but the nodes at the network boundary do not have any protection.

Many restoration mechanisms require a lot of signaling, so they usually do not meet real-time applications' requirement. The local restoration using ER-LSP proposed by [47] does not need fault notification and it does not need to wait for routing information synchronization. Although it needs to compute the ER-LSP to reroute, it does not give a burden to today's CPU. So it may be a fast restoration solution. However, the application scope of local restoration is limited.

Restoration mechanisms allocate resources after failure occurrence so they are resource effective but it takes time for them to provide recovery. Protection mechanisms provide fast recovery but they require additional resources. We should carefully consider the trade-off to choose the appropriate recovery mechanism so as to meet the requirements of users and network administration.

Compared to lower layer recovery mechanisms, the recovery mechanisms at the GMPLS level are relatively slow and may require more resources. Lower layer recovery mechanisms can provide fast recovery. But they have their limitations and disadvantages. For example, WDM networks may require complicated implementation and configuration for protection/restoration. And link layer protection cannot easily provide node protection.

In practice, usually a single type of protection mechanism does not satisfy the complicated working environment or user requirements. So a combination of recovery mechanisms is often the solution. When we choose a recovery solution, we need to achieve the balance between required resources and recovery time and the balance between cost and high survivability.

Nowadays, a lot of proposals have come up for LSP protection/restoration. GMPLS extends MPLS, but the LSP protection/restoration mechanisms that work in MPLS networks may not always work in GMPLS networks. For example, the "detour" proposal [48] makes the LSP very fault-tolerant in MPLS networks, but the current method described is only suitable for unidirectional LSPs. That is not applicable for GMPLS as bidirectional LSPs are recommended in GMPLS. Furthermore, the proposal places strict

constraints to the GMPLS network nodes when the "detour" LSP for protection is set up (see [51]). It is likely that the proposals that only work in MPLS networks but not in GMPLS networks would be dropped by IETF, e.g., [53] has been dropped, because of its limited scope.

Some proposals for LSP protection/restoration require the current signaling protocol to have more extensions, e.g., the one described in [48]. IETF considers these proposals very carefully as they would have a side-effect or put too much burden on the protocol. Some of these proposals are dropped, e.g., [54]. Therefore some researchers suggest that recovery mechanisms should be split from signaling protocol extensions (see [52]).

For local reroute, the aid from the signaling protocol is inevitable. But for the time being, none of the proposals in this area gets majority support. The issue is still being discussed in IETF.

With the further development of GMPLS, it is expected that more and more solutions are coming up for LSP protection/restoration in GMPLS.

## References:

- [1] E. Rosen, A. Viswanathan, et al., *Multiprotocol Label Switching Architecture*, RFC3031, IETF, <http://www.ietf.org>.
  
- [2] E. Mannie, et al., *Generalized Multi-Protocol Label Switching (GMPLS) Architecture*, draft-ietf-ccamp-gmpls-architecture-02.txt, work in progress, IETF, <http://www.ietf.org>.
  
- [3] P. Newman, G. Minshall, T. Lyon and L. Huston, *IP switching and gigabit routers*, IEEE communication magazines, January, 1997, pp.64-69.
  
- [4] P. Newman et al., *Ipsilon's General Switch Management Protocol Specification*, RFC1987, IETF, <http://www.ietf.org>.
  
- [5] P. Newman, W. L. Edwards, et al., *Transmission of Flow Labelled IPv4 on ATM Data Links*, RFC1954, IETF, <http://www.ietf.org>.
  
- [6] Y. Rekhter, B. Davie, et al., *Cisco Systems' Tag Switching Architecture Overview*, RFC2105, IETF, 1997.
  
- [7] C. Metz, *An overview of IP Switching Technology*, IBM Corporation, <http://www.networking.ibm.com/isr/ip/ipswp1.htm>.
  
- [8] Professor R. Jain, Department of Computer and Information Science, The Ohio State University, [http://www.cis.ohio-state.edu/~jain/cis788-97/ip\\_switching](http://www.cis.ohio-state.edu/~jain/cis788-97/ip_switching).
  
- [9] D. Awduche, Y. Rekhter, J. Drake, R. Coltun, *Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control With Optical Crossconnects*, draft-awduche-mpls-te-optical-03.txt, Work in Progress, April, 2001, IETF, <http://www.ietf.org>.
  
- [10] P. Ashwood-Smith et. al, *Generalized MPLS - Signaling Functional Description*, draft-ietf-mpls-generalized-signaling-02.txt, IETF Draft, Work in Progress, March, 2001, IETF, <http://www.ietf.org>.
  
- [11] E. Mannie et al., Section 3.2 of GMPLS Architecture, *draft-ietf-ccamp-gmpls-architecture-02.txt*, work in progress, IETF, <http://www.ietf.org>.
  
- [12] A. Banerjee, J. Drake, et al., *Generalized Multiprotocol Label Switching: An Overview of Signaling Enhancements and Recovery Techniques*. July 2001, IEEE Communication Magazine.
  
- [13] B. Rajagopalan, et al., *Abstract of IP over Optical Networks: A Framework*, draft-ietf-ipo-framework-01.txt, work in progress, IETF draft, <http://www.ietf.org>.



- [14] B. Rajagopalan, et al., *IP over Optical Networks: A Framework*, draft-ietf-ipo-framework-01.txt, IETF draft, <http://www.ietf.org>.
- [15] B. Rajagopalan, J. Luciani, et al., Section 3 of draft-many-ip-optical-framework-03.txt, work in progress, IETF.
- [16] D. Awduche, J. Malcolm, et al., Section 2 of *Requirements for Traffic Engineering Over MPLS* (RFC2702), IETF, <http://www.ietf.org>.
- [17] P. Srisuresh, P. Joseph, *TE LSAs to extend OSPF for Traffic Engineering*, draft-srisuresh-ospf-te-02.txt, work in progress, IETF, <http://www.ietf.org>.
- [18] D. Cheng, *OSPF Extensions to Support Multi-Area Traffic Engineering*, draft-cheng-ccamp-ospf-multiarea-te-extensions-00.txt, work in progress, IETF, <http://www.ietf.org>.
- [19] K. Kompella, Y. Rekhter, et al., *Routing Extensions in Support of Generalized MPLS*, draft-ietf-ccamp-gmpls-routing-04.txt, work in progress, IETF, <http://www.ietf.org>.
- [20] K. Kompella, Y. Rekhter, *LSP hierarchy with Generalized MPLS TE*, draft-ietf-mpls-lsp-hierarchy-06.txt, work in progress, IETF, <http://www.ietf.org>.
- [21] K. Kompella, Y. Rekhter, A. Banerjee, et al., *OSPF Extensions in Support of Generalized MPLS*, draft-ietf-ccamp-ospf-gmpls-extensions-07.txt, work in progress, IETF, <http://www.ietf.org>.
- [22] K. Kompella, Y. Rekhter, A. Banerjee, et al., *IS-IS Extensions in Support of Generalized MPLS*, draft-ietf-isis-gmpls-extensions-13.txt, work in progress, IETF, <http://www.ietf.org>.
- [23] B. Rajagopalan, J. Luciani, D. Awduche, et al., Section 8.2 of *IP over Optical Networks: A Framework*, draft-ietf-ipo-framework-01.txt, work in progress, IETF, <http://www.ietf.org>.
- [24] K. Kompella, Y. Rekhter, A. Banerjee, J. Drake, et al., *Routing Extensions in Support of Generalized MPLS*, draft-ietf-ccamp-gmpls-routing-04.txt, working in progress, IETF, <http://www.ietf.org>.
- [25] P. Ashwood-Smith, L. Berger, et al., *Generalized MPLS Signaling - CR-LDP Extensions*, draft-ietf-mpls-generalized-cr-ldp-06.txt, work in progress, IETF, <http://www.ietf.org>.

- [26] L. Berger, P. Ashwood-Smith, A. Banerjee, et al., *Generalized MPLS Signaling - RSVP-TE Extensions*, draft-ietf-mpls-generalized-rsvp-te-07.txt, work in progress, IETF, <http://www.ietf.org>.
- [27] L. Berger, P. Ashwood-Smith, A. Banerjee, G. Bernstein, et al., *Generalized MPLS - Signaling Functional Description*, draft-ietf-mpls-generalized-signaling-08.txt, work in progress, IETF, <http://www.ietf.org>.
- [28] L. Berger, P. Ashwood-Smith, A. Banerjee, G. Bernstein, et al., Section 4 of *Generalized MPLS - Signaling Functional Description*, draft-ietf-mpls-generalized-signaling-08.txt, work in progress, IETF, <http://www.ietf.org>.
- [29] B. Davie, Y. Rekhter, Section 2 of *MPLS Technology and Applications*, Morgan Kaufmann Publishers, 2000, ISBN 1558606564.
- [30] Y. Suemura, A. Kolarov, T. Shiragaki, *Protection of Hierarchical LSPs*, draft-suemura-protection-hierarchy-00.txt, work in progress, IETF, <http://www.ietf.org>.
- [31] V. Sharma, F. Hellstrand, et al., *Framework for MPLS-based Recovery*, Section 2, draft-ietf-mpls-recovery-frmwrk-04.txt, work in progress, IETF, <http://www.ietf.org>
- [32] B. Doshi, S. Dravida, P. Harshavardhana, O. Hauser, Y. Wang, *Optical Network Design and Restoration*, Bell Labs Technical Journal, Jan-March, 1999, see <http://www.lucent.com/minds/techjournal/pdf/jan-mar1999/paper04.pdf>
- [33] G. Maier, S. De Patre, M. Martinelli, et al., *Resilience schemes in WDM networks*, Italy, June 2001, see <http://leos.unipv.it/Pattavina.pdf>
- [34] V. Sharma, F. Hellstrand, et al., *Framework for MPLS-based Recovery*, Section 1.2, draft-ietf-mpls-recovery-frmwrk-04.txt, work in progress, IETF, <http://www.ietf.org>
- [35] W. S. Lai, D. McDysan, et al., Section 5.5, *Network Hierarchy and Multilayer Survivability*, draft-ietf-tewg-restore-hierarchy-00.txt, work in progress, IETF, <http://www.ietf.org>.
- [36] V. Sharma, B. Crane, et al., *Framework for MPLS-based Recovery*, draft-ietf-mpls-recovery-frmwrk-03.txt, work in progress, IETF, <http://www.ietf.org>.
- [37] D. Katz, D. Yeung, et al., *Traffic Engineering Extensions to OSPF*, draft-katz-yeung-ospf-traffic-06.txt, work in progress, IETF, <http://www.ietf.org>.
- [38] J. Lang, et al., *Link Management Protocol*, draft-ietf-ccamp-lmp-04.txt, work in progress, IETF, <http://www.ietf.org>.

- [39] L. Berger, P. Ashwood-Smith, A. Banerjee, et al., *Generalized MPLS Signaling - RSVP-TE Extensions*, draft-ietf-mpls-generalized-rsvp-te-07, work in progress, IETF, <http://www.ietf.org>.
- [40] L. Zhang, et al., Resource ReSerVation Protocol (RFC2205), IETF, <http://www.ietf.org>.
- [41] L. Andersson, P. Doolan, et al., *LDP Specification* (RFC3036), IETF, <http://www.ietf.org>.
- [42] B. Jamoussi, L. Andersson, et al., *Constraint-Based LSP Setup using LDP* (RFC3212), IETF, <http://www.ietf.org>.
- [43] P. Ashwood-Smith, L. Berger, et al., *Generalized MPLS Signaling - CR-LDP Extensions*, draft-ietf-mpls-generalized-cr-ldp-06.txt, working in progress, IETF, <http://www.ietf.org>.
- [44] A. Banerjee, J. Drake, et al., Section “GMPLS Protection and Restoration Techniques” of *Generalized Multiprotocol Label Switching: An Overview of Signaling Enhancements and Recovery Techniques*. July 2001, IEEE Communication Magazine.
- [45] V. Sharma, F. Hellstrand, et al., *Framework for MPLS-based Recovery*, Section 3.4, draft-ietf-mpls-recovery-frmwrk-04.txt, work in progress, IETF draft, 5, 2002.
- [46] Jeff Doyle, *Routing TCP/IP*, Volume 1, Cisco Press, 1998, ISBN 1578700418.
- [47] B. Davie, Y. Rekhter, Section 7 of *MPLS Technology and Applications*, Morgan Kaufmann Publishers, 2000, ISBN 1558606564.
- [48] Ping Pan, Der-Hwa Gan, et al., *Fast Reroute Extensions to RSVP-TE for LSP Tunnels*, draft-ietf-mpls-rsvp-lsp-fastreroute-00.txt, work in progress, July 2002, IETF, <http://www.ietf.org>.
- [49] Y. Suemura, A. Kolarov, T. Shiragaki, *Protection of Hierarchical LSPs*, draft-suemura-protection-hierarchy-00.txt, work in progress, IETF, <http://www.ietf.org>.
- [50] C. Ou, H. Zang, B. Mukherjee, *Sub-Path Protection for Scalability and Fast Recovery in WDM Mesh Networks*, Dept. of Computer Science, Univ. of California, Davis, CA, 2001.
- [51] B. Miller, E. Harrison, A. Farrel, *An examination of the methods for protecting MPLS LSPs against failures of network resources*, Data Connection, Oct 2001, <http://www.dataconnection.com/>.

[52] D. Papadimitriou, et al., *Restoration Mechanisms and Signaling in Optical Networks*, Proceedings of the Fiftieth Internet Engineering Task Force, March 18-23, 2001, <http://www.ietf.org/proceedings/01mar/slides/ccamp-7/>.

[53] C. Huang, V. Sharma, S. Makam and K. Owens, *A Path Protection/Restoration Mechanism for MPLS Networks*, draft-chang-mpls-path-protection-01.txt, work in progress, July, 2000, <http://www.ietf.org>.

[54] C. Huang, V. Sharma, S. Makam and K. Owens, *Extensions to RSVP-TE for MPLS Protection*, IETF, work in progress, IETF draft, June, 2000, <http://www.ietf.org>.

Name/Date
“BGP/MPLS VPNs,” RFC 2547 by E. Rosen et al., March 1999
“Framework for IP Based Virtual Private Networks,” RFC 2764 by B. Gleeson et al., February 2000
“LDP Specification,” RFC 3036 by L. Andersson et al., January 2001
“Multiprotocol Label Switching Architecture,” RFC 3031 by E. Rosen et al., January 2001
“TDM Service Specification for Pseudo-Wire Emulation Edge-to-Edge (PWE3),” version 3 by Prayson Pate et al., January 2001
“Ethernet Pseudo Wire Emulation Edge-to-Edge (PWE3),” version 0 by Tricci So et al., October 2001
“VPLS/LPE L2VPNs: Virtual Private LAN Services using Logical PE Architecture,” version 1 by Hamid Ould-Brahim et al., November 2001
“MPLS Concepts” by Cisco, 2002
“Using BGP as an Auto-Discovery Mechanism for Network-based VPNs,” version 2 by Hamid Ould-Brahim et al., January 2002
“Pseudo Wire (PW) Management Information Base,” version 2 by David Zelig et al., February 2002
“Pseudo Wire (PW) over MPLS PSN Management Information Base,” version 1 by David Zelig et al., February 2002
“Virtual Private LAN Service (VPLS) Solution Using GRE Based IP Tunnels,” version 0 by Tissa Senevirathne, February 2002
“Requirements for Virtual Private LAN Services (VPLS),” version 0 by Waldemar Augustyn, et al., March 2002
“Protocol Layering in PWE3,” version 0 by Stewart Bryant et al., May 2002
“SONET/SDH Circuit Emulation Service Over Packet (CEP) Management Information Base Using SMIPv2,” version 2 by Dave Danenberg et al., May 2002
“Discovering Nodes and Services in a VPLS Network,” version 0 by Olen Stokes et al., June 2002
“Framework for Pseudo Wire Emulation Edge-to-Edge (PWE3),” version 1 by Prayson Pate, ed., et al., June 2002
“Virtual Private LAN Services over MPLS,” version 2 by Marc Lasserre et al., June

Name/Date
2002
“BGP/MPLS VPNs,” version 2 by Eric. C. Rosen et al., July 2002
“TDM Service Specification for Pseudo-Wire Emulation Edge to Edge,” version 3 by Prayson Pate et al., August 2002
“Requirements for Virtual Private LAN Services (VPLS),” by Waldemar Augustyn, ed. et al., October 2002
“The Must-Have Reference for IP and Next Generation Networking” by Anritsu Company, October 23, 2002
“Ethernet Pseudo-wire over L2TPv3 (multipoint support),” version 0 by CY Lee and M. Higashiyama, November 2002
“Virtual Private LAN Service,” version 1 by K. Kompella et al., November 2002
“Ethernet Services—Service Definition and Market Potential” by Peter P. Komisarczuk, October - December 2002
“VPLS and BGP-based VPNs” by Jean-March Uzé, February 19, 2003
“PWE3 Architecture,” version 7 by Stewart Bryant and Prayson Pate, eds., March 2003
“Layer 2 VPNs Over Tunnels,” version 3 by K. Kompella et al., April 2003
“Pseudo Wire (PW) OAM Message Mapping,” version 0 by Thomas D. Nadeau and Monique Morrow, April 2003
“Pseudo Wire (PW) Virtual Circuit Connection Verification (VCCV),” version 0 by Thomas D. Nadeau et al., April 2003
“Pseudo Wire (PW) Virtual Circuit Connection Verification (VCCV),” version 1 by Thomas D. Nadeau and Rahul Aggarwal, June 2003
“Pseudowires and L2TPv3,” version 1 by W. M. Townsley, June 2003
“Pseudo Wire (PW) Virtual Circuit Connection Verification (VCCV),” version 0 by Thomas D. Nadeau and Rahul Aggarwal, July 2003
“IP-Only LAN Service (IPLS),” version 0 by Himanshu Shah et al., November 2003
“Layer 2 VPN experiences over a metro IPoDWDM network” by F. Valera et al., 2004
“Radius/L2TP Based VPLS,” version 0 by Juha Heinanen et al., January 2004

Name/Date
“Virtual Private LAN Service,” version 1 by K. Kompella, January 2004
“Virtual Private LAN Services over MPLS,” version 3 by Marc Lasserre and Vach Kompella, eds., April 2004
“BGP: The Next Best Thing Since Sliced Bread?” by Petri Miettinen, April 26-27, 2004
“Layer 2 and 3 Virtual Private Networks: Taxonomy, Technology, and Standardization Efforts” by Paul Knight et al., June 2004
“MPLS Basics and In-Depth: Overview of MPLS Fundamentals, Basic Operation, and In-Depth overview of Service Capabilities” by Craig Hill, June 29, 2004
“Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3),” RFC 3916 by X. Xiao et al., September 2004
“IANA Allocations for pseudo Wire Edge to Edge Emulation (PWE3),” version 7 by Luca Martini and W. Mark Townsley, October 2004
“Multicast in BGP/MPLS VPNs and VPLS,” version 1 by Rahul Aggarwal, ed., et al., October 2004
“Testing Hierarchical Virtual Private LAN Services,” version 0 by Olen Stokes et al., October 2004
“Pseudowire Setup and Maintenance using LDP,” version 14 by Luca Martini, ed., et al., December 2004
“Overview of the Internet Multicast Routing Architecture,” version 1 by P. Savola, December 20, 2004
“IANA Allocations for pseudo Wire Edge to Edge Emulation (PWE3),” version 8 by Luca Martini and W. Mark Townsley, February 2005
“Pseudo Wire Protection,” version 0 by Ping Pan, February 2005
“Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture,” RFC 3985 by S. Bryant et al., March 2005
“Pseudowire Setup and Maintenance using LDP,” version 16 by Luca Martini, ed., et al., March 2005
“VPLS Applicability,” version 0 by Marc Lasserre et al., March 2005
“Pseudo Wire Switching,” version 3 by Luca Martini et al., April 2005
“Virtual Private LAN Service,” version 5 by K. Kompella, April 8, 2005

Name/Date
“Circuit Cross-Connect,” version 2 by K. Kompella et al., April 15, 2005
“Atom Feed Autodiscovery,” version 1 by M. Pilgrim, May 10, 2005
“Constrained VPN Route Distribution,” version 2 by P. Marques et al., June 22, 2005
“Multicast in VPLS,” version 1 by R. Aggarwal and Y. Kamite, July 2005
“Propagation of VPLS IP Multicast Group Membership Information,” version 0 by R. Aggarwal and Y. Kamite, July 2005
“Supporting IP Multicast over VPLS,” version 3 by Y. Serbest et al., July 2005
“Label Distribution Protocol Extensions for Point-to-Multipoint Label Switched Paths,” version 1 by I. Minei et al., July 17, 2005
“Provisioning, Autodiscovery, and Signaling in L2VPNs,” version 6 by E. Rosen, et al., September 9, 2005
“Requirements for Multicast Support in Virtual Private LAN Services,” version 1 by Y. Kamite et al., September 15, 2005
“PWE3 Control Word for use over an MPLS PSN,” version 6 by S. Bryant et al., October 2005
“IANA Allocations for pseudo Wire Edge to Edge Emulation (PWE3),” version 15 by Luca Martini, November 2005
“PWE3 Fragmentation and Reassembly,” version 10 by Andrew G. Malis et al., November 2005
“Virtual Private LAN Services over MPLS,” version 8 by Marc Lasserre and Vach Kompella, eds., November 2005
“Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN,” RFC 4385 by S. Bryant et al., February 2006
“Telecommunications and Internet Converged Services and Protocols for Advanced Networking (TISPAN); NGN management; Operations Support Systems Architecture” by ETSI, March 2006
“Any Transport over MPLS by Cisco,” Multiple Cisco IOS Releases by Cisco
“Any Transport over MPLS,” Cisco IOS Release 12.0(26)S by Cisco
“Ten Features of NX-OS Every Customer Should Consider” by Ron Fuller



Network Working Group  
Request for Comments: 3209  
Category: Standards Track

D. Awduche  
Movaz Networks, Inc.  
L. Berger  
D. Gan  
Juniper Networks, Inc.  
T. Li  
Procket Networks, Inc.  
V. Srinivasan  
Cosine Communications, Inc.  
G. Swallow  
Cisco Systems, Inc.  
December 2001

## RSVP-TE: Extensions to RSVP for LSP Tunnels

### Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

### Copyright Notice

Copyright (C) The Internet Society (2001). All Rights Reserved.

### Abstract

This document describes the use of RSVP (Resource Reservation Protocol), including all the necessary extensions, to establish label-switched paths (LSPs) in MPLS (Multi-Protocol Label Switching). Since the flow along an LSP is completely identified by the label applied at the ingress node of the path, these paths may be treated as tunnels. A key application of LSP tunnels is traffic engineering with MPLS as specified in RFC 2702.

We propose several additional objects that extend RSVP, allowing the establishment of explicitly routed label switched paths using RSVP as a signaling protocol. The result is the instantiation of label-switched tunnels which can be automatically routed away from network failures, congestion, and bottlenecks.

## Contents

1	Introduction	3
1.1	Background	4
1.2	Terminology	6
2	Overview	7
2.1	LSP Tunnels and Traffic Engineered Tunnels	7
2.2	Operation of LSP Tunnels	8
2.3	Service Classes	10
2.4	Reservation Styles	10
2.4.1	Fixed Filter (FF) Style	10
2.4.2	Wildcard Filter (WF) Style	11
2.4.3	Shared Explicit (SE) Style	11
2.5	Rerouting Traffic Engineered Tunnels	12
2.6	Path MTU	13
3	LSP Tunnel related Message Formats	15
3.1	Path Message	15
3.2	Resv Message	16
4	LSP Tunnel related Objects	17
4.1	Label Object	17
4.1.1	Handling Label Objects in Resv messages	17
4.1.2	Non-support of the Label Object	19
4.2	Label Request Object	19
4.2.1	Label Request without Label Range	19
4.2.2	Label Request with ATM Label Range	20
4.2.3	Label Request with Frame Relay Label Range	21
4.2.4	Handling of LABEL_REQUEST	22
4.2.5	Non-support of the Label Request Object	23
4.3	Explicit Route Object	23
4.3.1	Applicability	24
4.3.2	Semantics of the Explicit Route Object	24
4.3.3	Subobjects	25
4.3.4	Processing of the Explicit Route Object	28
4.3.5	Loops	30
4.3.6	Forward Compatibility	30
4.3.7	Non-support of the Explicit Route Object	31
4.4	Record Route Object	31
4.4.1	Subobjects	31
4.4.2	Applicability	34
4.4.3	Processing RRO	35
4.4.4	Loop Detection	36
4.4.5	Forward Compatibility	37
4.4.6	Non-support of RRO	37
4.5	Error Codes for ERO and RRO	37
4.6	Session, Sender Template, and Filter Spec Objects	38
4.6.1	Session Object	39
4.6.2	Sender Template Object	40
4.6.3	Filter Specification Object	42

4.6.4	Reroute and Bandwidth Increase Procedure .....	42
4.7	Session Attribute Object .....	43
4.7.1	Format without resource affinities .....	43
4.7.2	Format with resource affinities .....	45
4.7.3	Procedures applying to both C-Types .....	46
4.7.4	Resource Affinity Procedures .....	48
5	Hello Extension .....	49
5.1	Hello Message Format .....	50
5.2	HELLO Object formats .....	51
5.2.1	HELLO REQUEST object .....	51
5.2.2	HELLO ACK object .....	51
5.3	Hello Message Usage .....	52
5.4	Multi-Link Considerations .....	53
5.5	Compatibility .....	54
6	Security Considerations .....	54
7	IANA Considerations .....	54
7.1	Message Types .....	55
7.2	Class Numbers and C-Types .....	55
7.3	Error Codes and Globally-Defined Error Value Sub-Codes .	57
7.4	Subobject Definitions .....	57
8	Intellectual Property Considerations .....	58
9	Acknowledgments .....	58
10	References .....	58
11	Authors' Addresses .....	60
12	Full Copyright Statement .....	61

## 1. Introduction

Section 2.9 of the MPLS architecture [2] defines a label distribution protocol as a set of procedures by which one Label Switched Router (LSR) informs another of the meaning of labels used to forward traffic between and through them. The MPLS architecture does not assume a single label distribution protocol. This document is a specification of extensions to RSVP for establishing label switched paths (LSPs) in MPLS networks.

Several of the new features described in this document were motivated by the requirements for traffic engineering over MPLS (see [3]). In particular, the extended RSVP protocol supports the instantiation of explicitly routed LSPs, with or without resource reservations. It also supports smooth rerouting of LSPs, preemption, and loop detection.

The LSPs created with RSVP can be used to carry the "Traffic Trunks" described in [3]. The LSP which carries a traffic trunk and a traffic trunk are distinct though closely related concepts. For example, two LSPs between the same source and destination could be load shared to carry a single traffic trunk. Conversely several

traffic trunks could be carried in the same LSP if, for instance, the LSP were capable of carrying several service classes. The applicability of these extensions is discussed further in [10].

Since the traffic that flows along a label-switched path is defined by the label applied at the ingress node of the LSP, these paths can be treated as tunnels, tunneling below normal IP routing and filtering mechanisms. When an LSP is used in this way we refer to it as an LSP tunnel.

LSP tunnels allow the implementation of a variety of policies related to network performance optimization. For example, LSP tunnels can be automatically or manually routed away from network failures, congestion, and bottlenecks. Furthermore, multiple parallel LSP tunnels can be established between two nodes, and traffic between the two nodes can be mapped onto the LSP tunnels according to local policy. Although traffic engineering (that is, performance optimization of operational networks) is expected to be an important application of this specification, the extended RSVP protocol can be used in a much wider context.

The purpose of this document is to describe the use of RSVP to establish LSP tunnels. The intent is to fully describe all the objects, packet formats, and procedures required to realize interoperable implementations. A few new objects are also defined that enhance management and diagnostics of LSP tunnels.

The document also describes a means of rapid node failure detection via a new HELLO message.

All objects and messages described in this specification are optional with respect to RSVP. This document discusses what happens when an object described here is not supported by a node.

Throughout this document, the discussion will be restricted to unicast label switched paths. Multicast LSPs are left for further study.

### 1.1. Background

Hosts and routers that support both RSVP [1] and Multi-Protocol Label Switching [2] can associate labels with RSVP flows. When MPLS and RSVP are combined, the definition of a flow can be made more flexible. Once a label switched path (LSP) is established, the traffic through the path is defined by the label applied at the ingress node of the LSP. The mapping of label to traffic can be accomplished using a number of different criteria. The set of packets that are assigned the same label value by a specific node are

said to belong to the same forwarding equivalence class (FEC) (see [2]), and effectively define the "RSVP flow." When traffic is mapped onto a label-switched path in this way, we call the LSP an "LSP Tunnel". When labels are associated with traffic flows, it becomes possible for a router to identify the appropriate reservation state for a packet based on the packet's label value.

The signaling protocol model uses downstream-on-demand label distribution. A request to bind labels to a specific LSP tunnel is initiated by an ingress node through the RSVP Path message. For this purpose, the RSVP Path message is augmented with a LABEL\_REQUEST object. Labels are allocated downstream and distributed (propagated upstream) by means of the RSVP Resv message. For this purpose, the RSVP Resv message is extended with a special LABEL object. The procedures for label allocation, distribution, binding, and stacking are described in subsequent sections of this document.

The signaling protocol model also supports explicit routing capability. This is accomplished by incorporating a simple EXPLICIT\_ROUTE object into RSVP Path messages. The EXPLICIT\_ROUTE object encapsulates a concatenation of hops which constitutes the explicitly routed path. Using this object, the paths taken by label-switched RSVP-MPLS flows can be pre-determined, independent of conventional IP routing. The explicitly routed path can be administratively specified, or automatically computed by a suitable entity based on QoS and policy requirements, taking into consideration the prevailing network state. In general, path computation can be control-driven or data-driven. The mechanisms, processes, and algorithms used to compute explicitly routed paths are beyond the scope of this specification.

One useful application of explicit routing is traffic engineering. Using explicitly routed LSPs, a node at the ingress edge of an MPLS domain can control the path through which traffic traverses from itself, through the MPLS network, to an egress node. Explicit routing can be used to optimize the utilization of network resources and enhance traffic oriented performance characteristics.

The concept of explicitly routed label switched paths can be generalized through the notion of abstract nodes. An abstract node is a group of nodes whose internal topology is opaque to the ingress node of the LSP. An abstract node is said to be simple if it contains only one physical node. Using this concept of abstraction, an explicitly routed LSP can be specified as a sequence of IP prefixes or a sequence of Autonomous Systems.

The signaling protocol model supports the specification of an explicit path as a sequence of strict and loose routes. The combination of abstract nodes, and strict and loose routes significantly enhances the flexibility of path definitions.

An advantage of using RSVP to establish LSP tunnels is that it enables the allocation of resources along the path. For example, bandwidth can be allocated to an LSP tunnel using standard RSVP reservations and Integrated Services service classes [4].

While resource reservations are useful, they are not mandatory. Indeed, an LSP can be instantiated without any resource reservations whatsoever. Such LSPs without resource reservations can be used, for example, to carry best effort traffic. They can also be used in many other contexts, including implementation of fall-back and recovery policies under fault conditions, and so forth.

## 1.2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC2119 [6].

The reader is assumed to be familiar with the terminology in [1], [2] and [3].

### Abstract Node

A group of nodes whose internal topology is opaque to the ingress node of the LSP. An abstract node is said to be simple if it contains only one physical node.

### Explicitly Routed LSP

An LSP whose path is established by a means other than normal IP routing.

### Label Switched Path

The path created by the concatenation of one or more label switched hops, allowing a packet to be forwarded by swapping labels from an MPLS node to another MPLS node. For a more precise definition see [2].

### LSP

A Label Switched Path

### LSP Tunnel

An LSP which is used to tunnel below normal IP routing and/or filtering mechanisms.

### Traffic Engineered Tunnel (TE Tunnel)

A set of one or more LSP Tunnels which carries a traffic trunk.

### Traffic Trunk

A set of flows aggregated by their service class and then placed on an LSP or set of LSPs called a traffic engineered tunnel. For further discussion see [3].

## 2. Overview

### 2.1. LSP Tunnels and Traffic Engineered Tunnels

According to [1], "RSVP defines a 'session' to be a data flow with a particular destination and transport-layer protocol." However, when RSVP and MPLS are combined, a flow or session can be defined with greater flexibility and generality. The ingress node of an LSP can use a variety of means to determine which packets are assigned a particular label. Once a label is assigned to a set of packets, the label effectively defines the "flow" through the LSP. We refer to such an LSP as an "LSP tunnel" because the traffic through it is opaque to intermediate nodes along the label switched path.

New RSVP SESSION, SENDER\_TEMPLATE, and FILTER\_SPEC objects, called LSP\_TUNNEL\_IPv4 and LSP\_TUNNEL\_IPv6 have been defined to support the LSP tunnel feature. The semantics of these objects, from the perspective of a node along the label switched path, is that traffic belonging to the LSP tunnel is identified solely on the basis of packets arriving from the PHOP or "previous hop" (see [1]) with the particular label value(s) assigned by this node to upstream senders to the session. In fact, the IPv4(v6) that appears in the object name only denotes that the destination address is an IPv4(v6) address. When we refer to these objects generically, we use the qualifier LSP\_TUNNEL.

In some applications it is useful to associate sets of LSP tunnels. This can be useful during reroute operations or to spread a traffic trunk over multiple paths. In the traffic engineering application such sets are called traffic engineered tunnels (TE tunnels). To enable the identification and association of such LSP tunnels, two identifiers are carried. A tunnel ID is part of the SESSION object. The SESSION object uniquely defines a traffic engineered tunnel. The

SENDER\_TEMPLATE and FILTER\_SPEC objects carry an LSP ID. The SENDER\_TEMPLATE (or FILTER\_SPEC) object together with the SESSION object uniquely identifies an LSP tunnel

## 2.2. Operation of LSP Tunnels

This section summarizes some of the features supported by RSVP as extended by this document related to the operation of LSP tunnels. These include: (1) the capability to establish LSP tunnels with or without QoS requirements, (2) the capability to dynamically reroute an established LSP tunnel, (3) the capability to observe the actual route traversed by an established LSP tunnel, (4) the capability to identify and diagnose LSP tunnels, (5) the capability to preempt an established LSP tunnel under administrative policy control, and (6) the capability to perform downstream-on-demand label allocation, distribution, and binding. In the following paragraphs, these features are briefly described. More detailed descriptions can be found in subsequent sections of this document.

To create an LSP tunnel, the first MPLS node on the path -- that is, the sender node with respect to the path -- creates an RSVP Path message with a session type of LSP\_TUNNEL\_IPv4 or LSP\_TUNNEL\_IPv6 and inserts a LABEL\_REQUEST object into the Path message. The LABEL\_REQUEST object indicates that a label binding for this path is requested and also provides an indication of the network layer protocol that is to be carried over this path. The reason for this is that the network layer protocol sent down an LSP cannot be assumed to be IP and cannot be deduced from the L2 header, which simply identifies the higher layer protocol as MPLS.

If the sender node has knowledge of a route that has high likelihood of meeting the tunnel's QoS requirements, or that makes efficient use of network resources, or that satisfies some policy criteria, the node can decide to use the route for some or all of its sessions. To do this, the sender node adds an EXPLICIT\_ROUTE object to the RSVP Path message. The EXPLICIT\_ROUTE object specifies the route as a sequence of abstract nodes.

If, after a session has been successfully established, the sender node discovers a better route, the sender can dynamically reroute the session by simply changing the EXPLICIT\_ROUTE object. If problems are encountered with an EXPLICIT\_ROUTE object, either because it causes a routing loop or because some intermediate routers do not support it, the sender node is notified.

By adding a RECORD\_ROUTE object to the Path message, the sender node can receive information about the actual route that the LSP tunnel traverses. The sender node can also use this object to request



notification from the network concerning changes to the routing path. The RECORD\_ROUTE object is analogous to a path vector, and hence can be used for loop detection.

Finally, a SESSION\_ATTRIBUTE object can be added to Path messages to aid in session identification and diagnostics. Additional control information, such as setup and hold priorities, resource affinities (see [3]), and local-protection, are also included in this object.

Routers along the path may use the setup and hold priorities along with SENDER\_TSPEC and any POLICY\_DATA objects contained in Path messages as input to policy control. For instance, in the traffic engineering application, it is very useful to use the Path message as a means of verifying that bandwidth exists at a particular priority along an entire path before preempting any lower priority reservations. If a Path message is allowed to progress when there are insufficient resources, then there is a danger that lower priority reservations downstream of this point will unnecessarily be preempted in a futile attempt to service this request.

When the EXPLICIT\_ROUTE object (ERO) is present, the Path message is forwarded towards its destination along a path specified by the ERO. Each node along the path records the ERO in its path state block. Nodes may also modify the ERO before forwarding the Path message. In this case the modified ERO SHOULD be stored in the path state block in addition to the received ERO.

The LABEL\_REQUEST object requests intermediate routers and receiver nodes to provide a label binding for the session. If a node is incapable of providing a label binding, it sends a PathErr message with an "unknown object class" error. If the LABEL\_REQUEST object is not supported end to end, the sender node will be notified by the first node which does not provide this support.

The destination node of a label-switched path responds to a LABEL\_REQUEST by including a LABEL object in its response RSVP Resv message. The LABEL object is inserted in the filter spec list immediately following the filter spec to which it pertains.

The Resv message is sent back upstream towards the sender, following the path state created by the Path message, in reverse order. Note that if the path state was created by use of an ERO, then the Resv message will follow the reverse path of the ERO.

Each node that receives a Resv message containing a LABEL object uses that label for outgoing traffic associated with this LSP tunnel. If the node is not the sender, it allocates a new label and places that label in the corresponding LABEL object of the Resv message which it

sends upstream to the PHOP. The label sent upstream in the LABEL object is the label which this node will use to identify incoming traffic associated with this LSP tunnel. This label also serves as shorthand for the Filter Spec. The node can now update its "Incoming Label Map" (ILM), which is used to map incoming labeled packets to a "Next Hop Label Forwarding Entry" (NHLFE), see [2].

When the Resv message propagates upstream to the sender node, a label-switched path is effectively established.

### 2.3. Service Classes

This document does not restrict the type of Integrated Service requests for reservations. However, an implementation SHOULD support the Controlled-Load service [4] and the Null Service [16].

### 2.4. Reservation Styles

The receiver node can select from among a set of possible reservation styles for each session, and each RSVP session must have a particular style. Senders have no influence on the choice of reservation style. The receiver can choose different reservation styles for different LSPs.

An RSVP session can result in one or more LSPs, depending on the reservation style chosen.

Some reservation styles, such as FF, dedicate a particular reservation to an individual sender node. Other reservation styles, such as WF and SE, can share a reservation among several sender nodes. The following sections discuss the different reservation styles and their advantages and disadvantages. A more detailed discussion of reservation styles can be found in [1].

#### 2.4.1. Fixed Filter (FF) Style

The Fixed Filter (FF) reservation style creates a distinct reservation for traffic from each sender that is not shared by other senders. This style is common for applications in which traffic from each sender is likely to be concurrent and independent. The total amount of reserved bandwidth on a link for sessions using FF is the sum of the reservations for the individual senders.

Because each sender has its own reservation, a unique label is assigned to each sender. This can result in a point-to-point LSP between every sender/receiver pair.

#### 2.4.2. Wildcard Filter (WF) Style

With the Wildcard Filter (WF) reservation style, a single shared reservation is used for all senders to a session. The total reservation on a link remains the same regardless of the number of senders.

A single multipoint-to-point label-switched-path is created for all senders to the session. On links that senders to the session share, a single label value is allocated to the session. If there is only one sender, the LSP looks like a normal point-to-point connection. When multiple senders are present, a multipoint-to-point LSP (a reversed tree) is created.

This style is useful for applications in which not all senders send traffic at the same time. A phone conference, for example, is an application where not all speakers talk at the same time. If, however, all senders send simultaneously, then there is no means of getting the proper reservations made. Either the reserved bandwidth on links close to the destination will be less than what is required or then the reserved bandwidth on links close to some senders will be greater than what is required. This restricts the applicability of WF for traffic engineering purposes.

Furthermore, because of the merging rules of WF, EXPLICIT\_ROUTE objects cannot be used with WF reservations. As a result of this issue and the lack of applicability to traffic engineering, use of WF is not considered in this document.

#### 2.4.3. Shared Explicit (SE) Style

The Shared Explicit (SE) style allows a receiver to explicitly specify the senders to be included in a reservation. There is a single reservation on a link for all the senders listed. Because each sender is explicitly listed in the Resv message, different labels may be assigned to different senders, thereby creating separate LSPs.

SE style reservations can be provided using multipoint-to-point label-switched-path or LSP per sender. Multipoint-to-point LSPs may be used when path messages do not carry the EXPLICIT\_ROUTE object, or when Path messages have identical EXPLICIT\_ROUTE objects. In either of these cases a common label may be assigned.

Path messages from different senders can each carry their own ERO, and the paths taken by the senders can converge and diverge at any point in the network topology. When Path messages have differing EXPLICIT\_ROUTE objects, separate LSPs for each EXPLICIT\_ROUTE object must be established.

## 2.5. Rerouting Traffic Engineered Tunnels

One of the requirements for Traffic Engineering is the capability to reroute an established TE tunnel under a number of conditions, based on administrative policy. For example, in some contexts, an administrative policy may dictate that a given TE tunnel is to be rerouted when a more "optimal" route becomes available. Another important context when TE tunnel reroute is usually required is upon failure of a resource along the TE tunnel's established path. Under some policies, it may also be necessary to return the TE tunnel to its original path when the failed resource becomes re-activated.

In general, it is highly desirable not to disrupt traffic, or adversely impact network operations while TE tunnel rerouting is in progress. This adaptive and smooth rerouting requirement necessitates establishing a new LSP tunnel and transferring traffic from the old LSP tunnel onto it before tearing down the old LSP tunnel. This concept is called "make-before-break." A problem can arise because the old and new LSP tunnels might compete with each other for resources on network segments which they have in common. Depending on availability of resources, this competition can cause Admission Control to prevent the new LSP tunnel from being established. An advantage of using RSVP to establish LSP tunnels is that it solves this problem very elegantly.

To support make-before-break in a smooth fashion, it is necessary that on links that are common to the old and new LSPs, resources used by the old LSP tunnel should not be released before traffic is transitioned to the new LSP tunnel, and reservations should not be counted twice because this might cause Admission Control to reject the new LSP tunnel.

A similar situation can arise when one wants to increase the bandwidth of a TE tunnel. The new reservation will be for the full amount needed, but the actual allocation needed is only the delta between the new and old bandwidth. If policy is being applied to PATH messages by intermediate nodes, then a PATH message requesting too much bandwidth will be rejected. In this situation simply increasing the bandwidth request without changing the SENDER\_TEMPLATE, could result in a tunnel being torn down, depending upon local policy.

The combination of the LSP\_TUNNEL SESSION object and the SE reservation style naturally accommodates smooth transitions in bandwidth and routing. The idea is that the old and new LSP tunnels share resources along links which they have in common. The LSP\_TUNNEL SESSION object is used to narrow the scope of the RSVP session to the particular TE tunnel in question. To uniquely identify a TE tunnel, we use the combination of the destination IP address (an address of the node which is the egress of the tunnel), a Tunnel ID, and the tunnel ingress node's IP address, which is placed in the Extended Tunnel ID field.

During the reroute or bandwidth-increase operation, the tunnel ingress needs to appear as two different senders to the RSVP session. This is achieved by the inclusion of the "LSP ID", which is carried in the SENDER\_TEMPLATE and FILTER\_SPEC objects. Since the semantics of these objects are changed, a new C-Types are assigned.

To effect a reroute, the ingress node picks a new LSP ID and forms a new SENDER\_TEMPLATE. The ingress node then creates a new ERO to define the new path. Thereafter the node sends a new Path Message using the original SESSION object and the new SENDER\_TEMPLATE and ERO. It continues to use the old LSP and refresh the old Path message. On links that are not held in common, the new Path message is treated as a conventional new LSP tunnel setup. On links held in common, the shared SESSION object and SE style allow the LSP to be established sharing resources with the old LSP. Once the ingress node receives a Resv message for the new LSP, it can transition traffic to it and tear down the old LSP.

To effect a bandwidth-increase, a new Path Message with a new LSP\_ID can be used to attempt a larger bandwidth reservation while the current LSP\_ID continues to be refreshed to ensure that the reservation is not lost if the larger reservation fails.

## 2.6. Path MTU

Standard RSVP [1] and Int-Serv [11] provide the RSVP sender with the minimum MTU available between the sender and the receiver. This path MTU identification capability is also provided for LSPs established via RSVP.

Path MTU information is carried, depending on which is present, in the Integrated Services or Null Service objects. When using Integrated Services objects, path MTU is provided based on the procedures defined in [11]. Path MTU identification when using Null Service objects is defined in [16].

With standard RSVP, the path MTU information is used by the sender to check which IP packets exceed the path MTU. For packets that exceed the path MTU, the sender either fragments the packets or, when the IP datagram has the "Don't Fragment" bit set, issues an ICMP destination unreachable message. This path MTU related handling is also required for LSPs established via RSVP.

The following algorithm applies to all unlabeled IP datagrams and to any labeled packets which the node knows to be IP datagrams, to which labels need to be added before forwarding. For labeled packets the bottom of stack is found, the IP header examined.

Using the terminology defined in [5], an LSR MUST execute the following algorithm:

1. Let N be the number of bytes in the label stack (i.e, 4 times the number of label stack entries) including labels to be added by this node.
2. Let M be the smaller of the "Maximum Initially Labeled IP Datagram Size" or of (Path MTU - N).

When the size of an IPv4 datagram (without labels) exceeds the value of M,

If the DF bit is not set in the IPv4 header, then

- (a) the datagram MUST be broken into fragments, each of whose size is no greater than M, and
- (b) each fragment MUST be labeled and then forwarded.

If the DF bit is set in the IPv4 header, then

- (a) the datagram MUST NOT be forwarded
- (b) Create an ICMP Destination Unreachable Message:
  - i. set its Code field [12] to "Fragmentation Required and DF Set",
  - ii. set its Next-Hop MTU field [13] to M
- (c) If possible, transmit the ICMP Destination Unreachable Message to the source of the of the discarded datagram.

When the size of an IPv6 datagram (without labels) exceeds the value of M,

- (a) the datagram MUST NOT be forwarded
- (b) Create an ICMP Packet too Big Message with the Next-Hop link MTU field [14] set to M
- (c) If possible, transmit the ICMP Packet too Big Message to the source of the of the discarded datagram.

### 3. LSP Tunnel related Message Formats

Five new objects are defined in this section:

Object name	Applicable RSVP messages
LABEL_REQUEST	Path
LABEL	Resv
EXPLICIT_ROUTE	Path
RECORD_ROUTE	Path, Resv
SESSION_ATTRIBUTE	Path

New C-Types are also assigned for the SESSION, SENDER\_TEMPLATE, and FILTER\_SPEC, objects.

Detailed descriptions of the new objects are given in later sections. All new objects are OPTIONAL with respect to RSVP. An implementation can choose to support a subset of objects. However, the LABEL\_REQUEST and LABEL objects are mandatory with respect to this specification.

The LABEL and RECORD\_ROUTE objects, are sender specific. In Resv messages they MUST appear after the associated FILTER\_SPEC and prior to any subsequent FILTER\_SPEC.

The relative placement of EXPLICIT\_ROUTE, LABEL\_REQUEST, and SESSION\_ATTRIBUTE objects is simply a recommendation. The ordering of these objects is not important, so an implementation MUST be prepared to accept objects in any order.

#### 3.1. Path Message

The format of the Path message is as follows:

```

<Path Message> ::=
    <Common Header> [ <INTEGRITY> ]
    <SESSION> <RSVP_HOP>
    <TIME_VALUES>
    [ <EXPLICIT_ROUTE> ]
    <LABEL_REQUEST>
    [ <SESSION_ATTRIBUTE> ]

```

```

    [ <POLICY_DATA> ... ]
    <sender descriptor>

```

```

<sender descriptor> ::= <SENDER_TEMPLATE> <SENDER_TSPEC>
    [ <ADSPEC> ]
    [ <RECORD_ROUTE> ]

```

### 3.2. Resv Message

The format of the Resv message is as follows:

```

<Resv Message> ::= <Common Header> [ <INTEGRITY> ]
    <SESSION> <RSVP_HOP>
    <TIME_VALUES>
    [ <RESV_CONFIRM> ] [ <SCOPE> ]
    [ <POLICY_DATA> ... ]
    <STYLE> <flow descriptor list>

<flow descriptor list> ::= <FF flow descriptor list>
    | <SE flow descriptor>

<FF flow descriptor list> ::= <FLOWSPEC> <FILTER_SPEC>
    <LABEL> [ <RECORD_ROUTE> ]
    | <FF flow descriptor list>
    <FF flow descriptor>

<FF flow descriptor> ::= [ <FLOWSPEC> ] <FILTER_SPEC> <LABEL>
    [ <RECORD_ROUTE> ]

<SE flow descriptor> ::= <FLOWSPEC> <SE filter spec list>

<SE filter spec list> ::= <SE filter spec>
    | <SE filter spec list> <SE filter spec>

<SE filter spec> ::= <FILTER_SPEC> <LABEL> [ <RECORD_ROUTE> ]

```

Note: LABEL and RECORD\_ROUTE (if present), are bound to the preceding FILTER\_SPEC. No more than one LABEL and/or RECORD\_ROUTE may follow each FILTER\_SPEC.



## 4. LSP Tunnel related Objects

### 4.1. Label Object

Labels MAY be carried in Resv messages. For the FF and SE styles, a label is associated with each sender. The label for a sender MUST immediately follow the FILTER\_SPEC for that sender in the Resv message.

The LABEL object has the following format:

LABEL class = 16, C\_Type = 1

```

      0                1                2                3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     (top label)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The contents of a LABEL is a single label, encoded in 4 octets. Each generic MPLS label is an unsigned integer in the range 0 through 1048575. Generic MPLS labels and FR labels are encoded right aligned in 4 octets. ATM labels are encoded with the VPI right justified in bits 0-15 and the VCI right justified in bits 16-31.

#### 4.1.1. Handling Label Objects in Resv messages

In MPLS a node may support multiple label spaces, perhaps associating a unique space with each incoming interface. For the purposes of the following discussion, the term "same label" means the identical label value drawn from the identical label space. Further, the following applies only to unicast sessions.

Labels received in Resv messages on different interfaces are always considered to be different even if the label value is the same.

##### 4.1.1.1. Downstream

The downstream node selects a label to represent the flow. If a label range has been specified in the label request, the label MUST be drawn from that range. If no label is available the node sends a PathErr message with an error code of "routing problem" and an error value of "label allocation failure".

If a node receives a Resv message that has assigned the same label value to multiple senders, then that node MAY also assign a single value to those same senders or to any subset of those senders. Note

that if a node intends to police individual senders to a session, it MUST assign unique labels to those senders.

In the case of ATM, one further condition applies. Some ATM nodes are not capable of merging streams. These nodes MAY indicate this by setting a bit in the label request to zero. The M-bit in the LABEL\_REQUEST object of C-Type 2, label request with ATM label range, serves this purpose. The M-bit SHOULD be set by nodes which are merge capable. If for any senders the M-bit is not set, the downstream node MUST assign unique labels to those senders.

Once a label is allocated, the node formats a new LABEL object. The node then sends the new LABEL object as part of the Resv message to the previous hop. The node SHOULD be prepared to forward packets carrying the assigned label prior to sending the Resv message. The LABEL object SHOULD be kept in the Reservation State Block. It is then used in the next Resv refresh event for formatting the Resv message.

A node is expected to send a Resv message before its refresh timers expire if the contents of the LABEL object change.

#### 4.1.1.2. Upstream

A node uses the label carried in the LABEL object as the outgoing label associated with the sender. The router allocates a new label and binds it to the incoming interface of this session/sender. This is the same interface that the router uses to forward Resv messages to the previous hops.

Several circumstance can lead to an unacceptable label.

1. the node is a merge incapable ATM switch but the downstream node has assigned the same label to two senders
2. The implicit null label was assigned, but the node is not capable of doing a penultimate pop for the associated L3PID
3. The assigned label is outside the requested label range

In any of these events the node send a ResvErr message with an error code of "routing problem" and an error value of "unacceptable label value".

4.1.2. Non-support of the Label Object

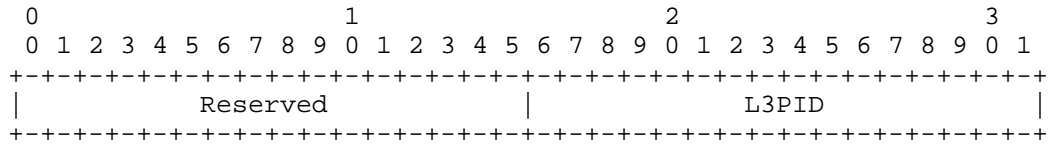
Under normal circumstances, a node should never receive a LABEL object in a Resv message unless it had included a LABEL\_REQUEST object in the corresponding Path message. However, an RSVP router that does not recognize the LABEL object sends a ResvErr with the error code "Unknown object class" toward the receiver. This causes the reservation to fail.

4.2. Label Request Object

The Label Request Class is 19. Currently there are three possible C\_Types. Type 1 is a Label Request without label range. Type 2 is a label request with an ATM label range. Type 3 is a label request with a Frame Relay label range. The LABEL\_REQUEST object formats are shown below.

4.2.1. Label Request without Label Range

Class = 19, C\_Type = 1



Reserved

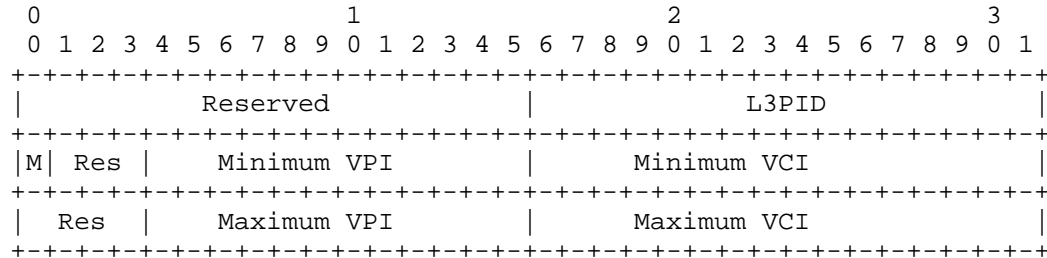
This field is reserved. It MUST be set to zero on transmission and MUST be ignored on receipt.

L3PID

an identifier of the layer 3 protocol using this path. Standard Ethertype values are used.

## 4.2.2. Label Request with ATM Label Range

Class = 19, C\_Type = 2



## Reserved (Res)

This field is reserved. It MUST be set to zero on transmission and MUST be ignored on receipt.

## L3PID

an identifier of the layer 3 protocol using this path. Standard Ethertype values are used.

## M

Setting this bit to one indicates that the node is capable of merging in the data plane

## Minimum VPI (12 bits)

This 12 bit field specifies the lower bound of a block of Virtual Path Identifiers that is supported on the originating switch. If the VPI is less than 12-bits it MUST be right justified in this field and preceding bits MUST be set to zero.

## Minimum VCI (16 bits)

This 16 bit field specifies the lower bound of a block of Virtual Connection Identifiers that is supported on the originating switch. If the VCI is less than 16-bits it MUST be right justified in this field and preceding bits MUST be set to zero.

Maximum VPI (12 bits)

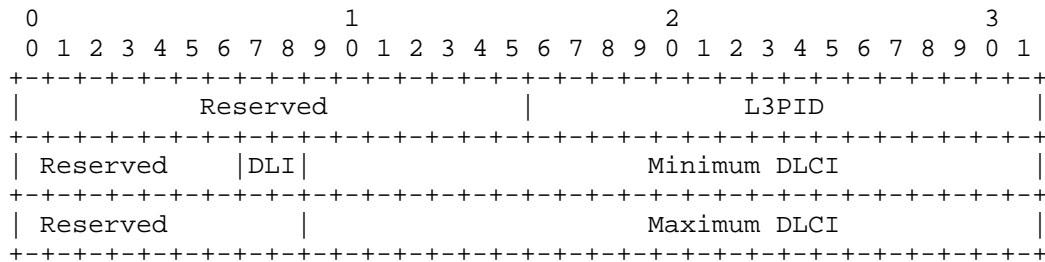
This 12 bit field specifies the upper bound of a block of Virtual Path Identifiers that is supported on the originating switch. If the VPI is less than 12-bits it MUST be right justified in this field and preceding bits MUST be set to zero.

Maximum VCI (16 bits)

This 16 bit field specifies the upper bound of a block of Virtual Connection Identifiers that is supported on the originating switch. If the VCI is less than 16-bits it MUST be right justified in this field and preceding bits MUST be set to zero.

4.2.3. Label Request with Frame Relay Label Range

Class = 19, C\_Type = 3



Reserved

This field is reserved. It MUST be set to zero on transmission and ignored on receipt.

L3PID

an identifier of the layer 3 protocol using this path. Standard Ethertype values are used.

DLI

DLCI Length Indicator. The number of bits in the DLCI. The following values are supported:

Len	DLCI bits
0	10
2	23

#### Minimum DLCI

This 23-bit field specifies the lower bound of a block of Data Link Connection Identifiers (DLCIs) that is supported on the originating switch. The DLCI MUST be right justified in this field and unused bits MUST be set to 0.

#### Maximum DLCI

This 23-bit field specifies the upper bound of a block of Data Link Connection Identifiers (DLCIs) that is supported on the originating switch. The DLCI MUST be right justified in this field and unused bits MUST be set to 0.

#### 4.2.4. Handling of LABEL\_REQUEST

To establish an LSP tunnel the sender creates a Path message with a LABEL\_REQUEST object. The LABEL\_REQUEST object indicates that a label binding for this path is requested and provides an indication of the network layer protocol that is to be carried over this path. This permits non-IP network layer protocols to be sent down an LSP. This information can also be useful in actual label allocation, because some reserved labels are protocol specific, see [5].

The LABEL\_REQUEST SHOULD be stored in the Path State Block, so that Path refresh messages will also contain the LABEL\_REQUEST object. When the Path message reaches the receiver, the presence of the LABEL\_REQUEST object triggers the receiver to allocate a label and to place the label in the LABEL object for the corresponding Resv message. If a label range was specified, the label MUST be allocated from that range. A receiver that accepts a LABEL\_REQUEST object MUST include a LABEL object in Resv messages pertaining to that Path message. If a LABEL\_REQUEST object was not present in the Path message, a node MUST NOT include a LABEL object in a Resv message for that Path message's session and PHOP.

A node that sends a LABEL\_REQUEST object MUST be ready to accept and correctly process a LABEL object in the corresponding Resv messages.

A node that recognizes a LABEL\_REQUEST object, but that is unable to support it (possibly because of a failure to allocate labels) SHOULD send a PathErr with the error code "Routing problem" and the error value "MPLS label allocation failure." This includes the case where a label range has been specified and a label cannot be allocated from that range.

A node which receives and forwards a Path message each with a LABEL\_REQUEST object, MUST copy the L3PID from the received LABEL\_REQUEST object to the forwarded LABEL\_REQUEST object.

If the receiver cannot support the protocol L3PID, it SHOULD send a PathErr with the error code "Routing problem" and the error value "Unsupported L3PID." This causes the RSVP session to fail.

#### 4.2.5. Non-support of the Label Request Object

An RSVP router that does not recognize the LABEL\_REQUEST object sends a PathErr with the error code "Unknown object class" toward the sender. An RSVP router that recognizes the LABEL\_REQUEST object but does not recognize the C\_Type sends a PathErr with the error code "Unknown object C\_Type" toward the sender. This causes the path setup to fail. The sender should notify management that a LSP cannot be established and possibly take action to continue the reservation without the LABEL\_REQUEST.

RSVP is designed to cope gracefully with non-RSVP routers anywhere between senders and receivers. However, obviously, non-RSVP routers cannot convey labels via RSVP. This means that if a router has a neighbor that is known to not be RSVP capable, the router MUST NOT advertise the LABEL\_REQUEST object when sending messages that pass through the non-RSVP routers. The router SHOULD send a PathErr back to the sender, with the error code "Routing problem" and the error value "MPLS being negotiated, but a non-RSVP capable router stands in the path." This same message SHOULD be sent, if a router receives a LABEL\_REQUEST object in a message from a non-RSVP capable router. See [1] for a description of how a downstream router can determine the presence of non-RSVP routers.

#### 4.3. Explicit Route Object

Explicit routes are specified via the EXPLICIT\_ROUTE object (ERO). The Explicit Route Class is 20. Currently one C\_Type is defined, Type 1 Explicit Route. The EXPLICIT\_ROUTE object has the following format:

Class = 20, C\_Type = 1

```

      0                1                2                3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|
|//                               (Subobjects)                               //|
|
+-----+-----+-----+-----+-----+-----+-----+-----+

```

#### Subobjects

The contents of an EXPLICIT\_ROUTE object are a series of variable-length data items called subobjects. The subobjects are defined in section 4.3.3 below.

If a Path message contains multiple EXPLICIT\_ROUTE objects, only the first object is meaningful. Subsequent EXPLICIT\_ROUTE objects MAY be ignored and SHOULD NOT be propagated.

#### 4.3.1. Applicability

The EXPLICIT\_ROUTE object is intended to be used only for unicast situations. Applications of explicit routing to multicast are a topic for further research.

The EXPLICIT\_ROUTE object is to be used only when all routers along the explicit route support RSVP and the EXPLICIT\_ROUTE object. The EXPLICIT\_ROUTE object is assigned a class value of the form 0bbbbbbb. RSVP routers that do not support the object will therefore respond with an "Unknown Object Class" error.

#### 4.3.2. Semantics of the Explicit Route Object

An explicit route is a particular path in the network topology. Typically, the explicit route is determined by a node, with the intent of directing traffic along that path.

An explicit route is described as a list of groups of nodes along the explicit route. In addition to the ability to identify specific nodes along the path, an explicit route can identify a group of nodes that must be traversed along the path. This capability allows the routing system a significant amount of local flexibility in fulfilling a request for an explicit route. This capability allows the generator of the explicit route to have imperfect information about the details of the path.



The explicit route is encoded as a series of subobjects contained in an EXPLICIT\_ROUTE object. Each subobject identifies a group of nodes in the explicit route. An explicit route is thus a specification of groups of nodes to be traversed.

To formalize the discussion, we call each group of nodes an abstract node. Thus, we say that an explicit route is a specification of a set of abstract nodes to be traversed. If an abstract node consists of only one node, we refer to it as a simple abstract node.

As an example of the concept of abstract nodes, consider an explicit route that consists solely of Autonomous System number subobjects. Each subobject corresponds to an Autonomous System in the global topology. In this case, each Autonomous System is an abstract node, and the explicit route is a path that includes each of the specified Autonomous Systems. There may be multiple hops within each Autonomous System, but these are opaque to the source node for the explicit route.

#### 4.3.3. Subobjects

The contents of an EXPLICIT\_ROUTE object are a series of variable-length data items called subobjects. Each subobject has the form:

```

0                               1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|L|   Type   |   Length   | (Subobject contents) |
+-----+-----+-----+-----+-----+-----+

```

L

The L bit is an attribute of the subobject. The L bit is set if the subobject represents a loose hop in the explicit route. If the bit is not set, the subobject represents a strict hop in the explicit route.

Type

The Type indicates the type of contents of the subobject. Currently defined values are:

- 1 IPv4 prefix
- 2 IPv6 prefix
- 32 Autonomous system number

### Length

The Length contains the total length of the subobject in bytes, including the L, Type and Length fields. The Length MUST be at least 4, and MUST be a multiple of 4.

#### 4.3.3.1. Strict and Loose Subobjects

The L bit in the subobject is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.' For brevity, we say that if the value of the subobject attribute is 'loose' then it is a 'loose subobject.' Otherwise, it's a 'strict subobject.' Further, we say that the abstract node of a strict or loose subobject is a strict or a loose node, respectively. Loose and strict nodes are always interpreted relative to their prior abstract nodes.

The path between a strict node and its preceding node MUST include only network nodes from the strict node and its preceding abstract node.

The path between a loose node and its preceding node MAY include other network nodes that are not part of the strict node or its preceding abstract node.

#### 4.3.3.2. Subobject 1: IPv4 prefix

0	1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9 0 1
+-----+			
L	Type	Length	IPv4 address (4 bytes)
+-----+			
	IPv4 address (continued)	Prefix Length	Resvd
+-----+			

### L

The L bit is an attribute of the subobject. The L bit is set if the subobject represents a loose hop in the explicit route. If the bit is not set, the subobject represents a strict hop in the explicit route.

### Type

0x01 IPv4 address

Length

The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length is always 8.

IPv4 address

An IPv4 address. This address is treated as a prefix based on the prefix length value below. Bits beyond the prefix are ignored on receipt and SHOULD be set to zero on transmission.

Prefix length

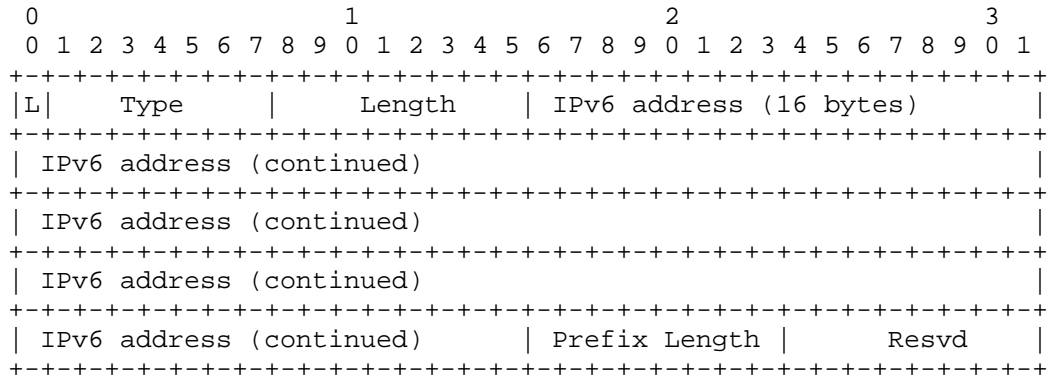
Length in bits of the IPv4 prefix

Padding

Zero on transmission. Ignored on receipt.

The contents of an IPv4 prefix subobject are a 4-octet IPv4 address, a 1-octet prefix length, and a 1-octet pad. The abstract node represented by this subobject is the set of nodes that have an IP address which lies within this prefix. Note that a prefix length of 32 indicates a single IPv4 node.

4.3.3.3. Subobject 2: IPv6 Prefix



L

The L bit is an attribute of the subobject. The L bit is set if the subobject represents a loose hop in the explicit route. If the bit is not set, the subobject represents a strict hop in the explicit route.

**Type**

0x02 IPv6 address

**Length**

The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length is always 20.

**IPv6 address**

An IPv6 address. This address is treated as a prefix based on the prefix length value below. Bits beyond the prefix are ignored on receipt and SHOULD be set to zero on transmission.

**Prefix Length**

Length in bits of the IPv6 prefix.

**Padding**

Zero on transmission. Ignored on receipt.

The contents of an IPv6 prefix subobject are a 16-octet IPv6 address, a 1-octet prefix length, and a 1-octet pad. The abstract node represented by this subobject is the set of nodes that have an IP address which lies within this prefix. Note that a prefix length of 128 indicates a single IPv6 node.

**4.3.3.4. Subobject 32: Autonomous System Number**

The contents of an Autonomous System (AS) number subobject are a 2-octet AS number. The abstract node represented by this subobject is the set of nodes belonging to the autonomous system.

The length of the AS number subobject is 4 octets.

**4.3.4. Processing of the Explicit Route Object****4.3.4.1. Selection of the Next Hop**

A node receiving a Path message containing an EXPLICIT\_ROUTE object must determine the next hop for this path. This is necessary because the next abstract node along the explicit route might be an IP subnet or an Autonomous System. Therefore, selection of this next hop may involve a decision from a set of feasible alternatives. The criteria used to make a selection from feasible alternatives is implementation dependent and can also be impacted by local policy, and is beyond the

scope of this specification. However, it is assumed that each node will make a best effort attempt to determine a loop-free path. Note that paths so determined can be overridden by local policy.

To determine the next hop for the path, a node performs the following steps:

- 1) The node receiving the RSVP message MUST first evaluate the first subobject. If the node is not part of the abstract node described by the first subobject, it has received the message in error and SHOULD return a "Bad initial subobject" error. If there is no first subobject, the message is also in error and the system SHOULD return a "Bad EXPLICIT\_ROUTE object" error.
- 2) If there is no second subobject, this indicates the end of the explicit route. The EXPLICIT\_ROUTE object SHOULD be removed from the Path message. This node may or may not be the end of the path. Processing continues with section 4.3.4.2, where a new EXPLICIT\_ROUTE object MAY be added to the Path message.
- 3) Next, the node evaluates the second subobject. If the node is also a part of the abstract node described by the second subobject, then the node deletes the first subobject and continues processing with step 2, above. Note that this makes the second subobject into the first subobject of the next iteration and allows the node to identify the next abstract node on the path of the message after possible repeated application(s) of steps 2 and 3.
- 4) Abstract Node Border Case: The node determines whether it is topologically adjacent to the abstract node described by the second subobject. If so, the node selects a particular next hop which is a member of the abstract node. The node then deletes the first subobject and continues processing with section 4.3.4.2.
- 5) Interior of the Abstract Node Case: Otherwise, the node selects a next hop within the abstract node of the first subobject (which the node belongs to) that is along the path to the abstract node of the second subobject (which is the next abstract node). If no such path exists then there are two cases:
  - 5a) If the second subobject is a strict subobject, there is an error and the node SHOULD return a "Bad strict node" error.
  - 5b) Otherwise, if the second subobject is a loose subobject, the node selects any next hop that is along the path to the next abstract node. If no path exists, there is an error, and the node SHOULD return a "Bad loose node" error.

- 6) Finally, the node replaces the first subobject with any subobject that denotes an abstract node containing the next hop. This is necessary so that when the explicit route is received by the next hop, it will be accepted.

#### 4.3.4.2. Adding subobjects to the Explicit Route Object

After selecting a next hop, the node MAY alter the explicit route in the following ways.

If, as part of executing the algorithm in section 4.3.4.1, the

EXPLICIT\_ROUTE object is removed, the node MAY add a new EXPLICIT\_ROUTE object.

Otherwise, if the node is a member of the abstract node for the first subobject, a series of subobjects MAY be inserted before the first subobject or MAY replace the first subobject. Each subobject in this series MUST denote an abstract node that is a subset of the current abstract node.

Alternately, if the first subobject is a loose subobject, an arbitrary series of subobjects MAY be inserted prior to the first subobject.

#### 4.3.5. Loops

While the EXPLICIT\_ROUTE object is of finite length, the existence of loose nodes implies that it is possible to construct forwarding loops during transients in the underlying routing protocol. This can be detected by the originator of the explicit route through the use of another opaque route object called the RECORD\_ROUTE object. The RECORD\_ROUTE object is used to collect detailed path information and is useful for loop detection and for diagnostics.

#### 4.3.6. Forward Compatibility

It is anticipated that new subobjects may be defined over time. A node which encounters an unrecognized subobject during its normal ERO processing sends a PathErr with the error code "Routing Error" and error value of "Bad Explicit Route Object" toward the sender. The EXPLICIT\_ROUTE object is included, truncated (on the left) to the offending subobject. The presence of an unrecognized subobject which is not encountered in a node's ERO processing SHOULD be ignored. It is passed forward along with the rest of the remaining ERO stack.

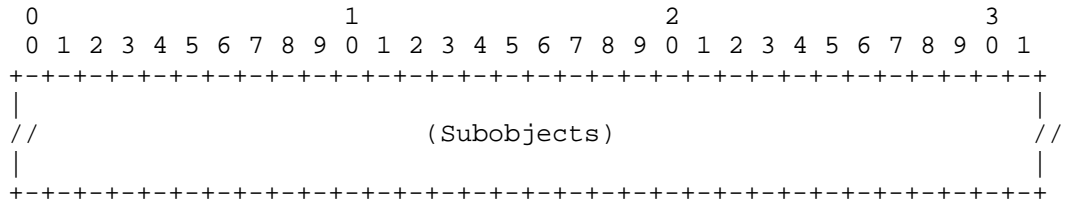
4.3.7. Non-support of the Explicit Route Object

An RSVP router that does not recognize the EXPLICIT\_ROUTE object sends a PathErr with the error code "Unknown object class" toward the sender. This causes the path setup to fail. The sender should notify management that a LSP cannot be established and possibly take action to continue the reservation without the EXPLICIT\_ROUTE or via a different explicit route.

4.4. Record Route Object

Routes can be recorded via the RECORD\_ROUTE object (RRO). Optionally, labels may also be recorded. The Record Route Class is 21. Currently one C\_Type is defined, Type 1 Record Route. The RECORD\_ROUTE object has the following format:

Class = 21, C\_Type = 1



Subobjects

The contents of a RECORD\_ROUTE object are a series of variable-length data items called subobjects. The subobjects are defined in section 4.4.1 below.

The RRO can be present in both RSVP Path and Resv messages. If a Path message contains multiple RROs, only the first RRO is meaningful. Subsequent RROs SHOULD be ignored and SHOULD NOT be propagated. Similarly, if in a Resv message multiple RROs are encountered following a FILTER\_SPEC before another FILTER\_SPEC is encountered, only the first RRO is meaningful. Subsequent RROs SHOULD be ignored and SHOULD NOT be propagated.

4.4.1. Subobjects

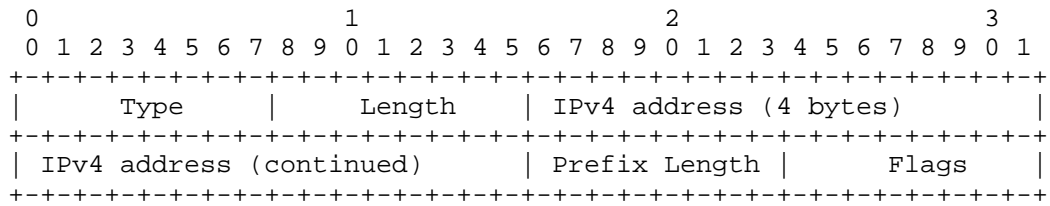
The contents of a RECORD\_ROUTE object are a series of variable-length data items called subobjects. Each subobject has its own Length field. The length contains the total length of the subobject in bytes, including the Type and Length fields. The length MUST always be a multiple of 4, and at least 4.

Subobjects are organized as a last-in-first-out stack. The first subobject relative to the beginning of RRO is considered the top. The last subobject is considered the bottom. When a new subobject is added, it is always added to the top.

An empty RRO with no subobjects is considered illegal.

Three kinds of subobjects are currently defined.

4.4.1.1. Subobject 1: IPv4 address



Type

0x01 IPv4 address

Length

The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length is always 8.

IPv4 address

A 32-bit unicast, host address. Any network-reachable interface address is allowed here. Illegal addresses, such as certain loopback addresses, SHOULD NOT be used.

Prefix length

32

Flags

0x01 Local protection available

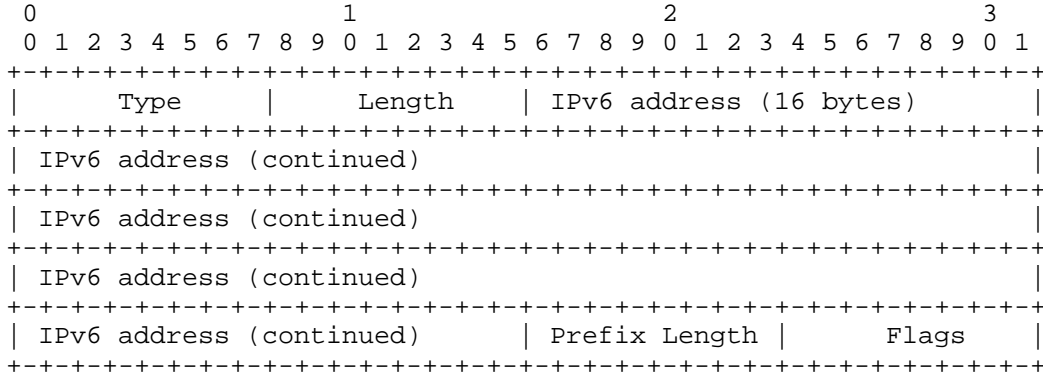
Indicates that the link downstream of this node is protected via a local repair mechanism. This flag can only be set if the Local protection flag was set in the SESSION\_ATTRIBUTE object of the corresponding Path message.



0x02 Local protection in use

Indicates that a local repair mechanism is in use to maintain this tunnel (usually in the face of an outage of the link it was previously routed over).

4.4.1.2. Subobject 2: IPv6 address



Type

0x02 IPv6 address

Length

The Length contains the total length of the subobject in bytes, including the Type and Length fields. The Length is always 20.

IPv6 address

A 128-bit unicast host address.

Prefix length

128

Flags

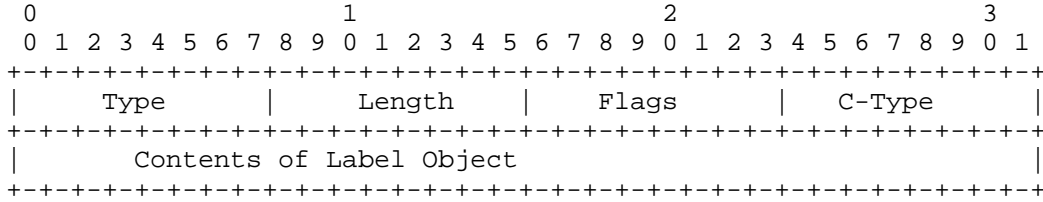
0x01 Local protection available

Indicates that the link downstream of this node is protected via a local repair mechanism. This flag can only be set if the Local protection flag was set in the SESSION\_ATTRIBUTE object of the corresponding Path message.

0x02 Local protection in use

Indicates that a local repair mechanism is in use to maintain this tunnel (usually in the face of an outage of the link it was previously routed over).

4.4.1.3. Subobject 3, Label



Type

0x03 Label

Length

The Length contains the total length of the subobject in bytes, including the Type and Length fields.

Flags

0x01 = Global label  
This flag indicates that the label will be understood if received on any interface.

C-Type

The C-Type of the included Label Object. Copied from the Label Object.

Contents of Label Object

The contents of the Label Object. Copied from the Label Object

4.4.2. Applicability

Only the procedures for use in unicast sessions are defined here.

There are three possible uses of RRO in RSVP. First, an RRO can function as a loop detection mechanism to discover L3 routing loops, or loops inherent in the explicit route. The exact procedure for doing so is described later in this document.

Second, an RRO collects up-to-date detailed path information hop-by-hop about RSVP sessions, providing valuable information to the sender or receiver. Any path change (due to network topology changes) will be reported.

Third, RRO syntax is designed so that, with minor changes, the whole object can be used as input to the EXPLICIT\_ROUTE object. This is useful if the sender receives RRO from the receiver in a Resv message, applies it to EXPLICIT\_ROUTE object in the next Path message in order to "pin down session path".

#### 4.4.3. Processing RRO

Typically, a node initiates an RSVP session by adding the RRO to the Path message. The initial RRO contains only one subobject - the sender's IP addresses. If the node also desires label recording, it sets the Label\_Recording flag in the SESSION\_ATTRIBUTE object.

When a Path message containing an RRO is received by an intermediate router, the router stores a copy of it in the Path State Block. The RRO is then used in the next Path refresh event for formatting Path messages. When a new Path message is to be sent, the router adds a new subobject to the RRO and appends the resulting RRO to the Path message before transmission.

The newly added subobject MUST be this router's IP address. The address to be added SHOULD be the interface address of the outgoing Path messages. If there are multiple addresses to choose from, the decision is a local matter. However, it is RECOMMENDED that the same address be chosen consistently.

When the Label\_Recording flag is set in the SESSION\_ATTRIBUTE object, nodes doing route recording SHOULD include a Label Record subobject. If the node is using a global label space, then it SHOULD set the Global Label flag.

The Label Record subobject is pushed onto the RECORD\_ROUTE object prior to pushing on the node's IP address. A node MUST NOT push on a Label Record subobject without also pushing on an IPv4 or IPv6 subobject.

Note that on receipt of the initial Path message, a node is unlikely to have a label to include. Once a label is obtained, the node SHOULD include the label in the RRO in the next Path refresh event.

If the newly added subobject causes the RRO to be too big to fit in a Path (or Resv) message, the RRO object SHALL be dropped from the message and message processing continues as normal. A PathErr (or

ResvErr) message SHOULD be sent back to the sender (or receiver). An error code of "Notify" and an error value of "RRO too large for MTU" is used. If the receiver receives such a ResvErr, it SHOULD send a PathErr message with error code of "Notify" and an error value of "RRO notification".

A sender receiving either of these error values SHOULD remove the RRO from the Path message.

Nodes SHOULD resend the above PathErr or ResvErr message each  $n$  seconds where  $n$  is the greater of 15 and the refresh interval for the associated Path or RESV message. The node MAY apply limits and/or back-off timers to limit the number of messages sent.

An RSVP router can decide to send Path messages before its refresh time if the RRO in the next Path message is different from the previous one. This can happen if the contents of the RRO received from the previous hop router changes or if this RRO is newly added to (or deleted from) the Path message.

When the destination node of an RSVP session receives a Path message with an RRO, this indicates that the sender node needs route recording. The destination node initiates the RRO process by adding an RRO to Resv messages. The processing mirrors that of the Path messages. The only difference is that the RRO in a Resv message records the path information in the reverse direction.

Note that each node along the path will now have the complete route from source to destination. The Path RRO will have the route from the source to this node; the Resv RRO will have the route from this node to the destination. This is useful for network management.

A received Path message without an RRO indicates that the sender node no longer needs route recording. Subsequent Resv messages SHALL NOT contain an RRO.

#### 4.4.4. Loop Detection

As part of processing an incoming RRO, an intermediate router looks into all subobjects contained within the RRO. If the router determines that it is already in the list, a forwarding loop exists.

An RSVP session is loop-free if downstream nodes receive Path messages or upstream nodes receive Resv messages with no routing loops detected in the contained RRO.

There are two broad classifications of forwarding loops. The first class is the transient loop, which occurs as a normal part of operations as L3 routing tries to converge on a consistent forwarding path for all destinations. The second class of forwarding loop is the permanent loop, which normally results from network mis-configuration.

The action performed by a node on receipt of an RRO depends on the message type in which the RRO is received.

For Path messages containing a forwarding loop, the router builds and sends a "Routing problem" PathErr message, with the error value "loop detected," and drops the Path message. Until the loop is eliminated, this session is not suitable for forwarding data packets. How the loop eliminated is beyond the scope of this document.

For Resv messages containing a forwarding loop, the router simply drops the message. Resv messages should not loop if Path messages do not loop.

#### 4.4.5. Forward Compatibility

New subobjects may be defined for the RRO. When processing an RRO, unrecognized subobjects SHOULD be ignored and passed on. When processing an RRO for loop detection, a node SHOULD parse over any unrecognized objects. Loop detection works by detecting subobjects which were inserted by the node itself on an earlier pass of the object. This ensures that the subobjects necessary for loop detection are always understood.

#### 4.4.6. Non-support of RRO

The RRO object is to be used only when all routers along the path support RSVP and the RRO object. The RRO object is assigned a class value of the form 0bbbbbbb. RSVP routers that do not support the object will therefore respond with an "Unknown Object Class" error.

#### 4.5. Error Codes for ERO and RRO

In the processing described above, certain errors must be reported as either a "Routing Problem" or "Notify". The value of the "Routing Problem" error code is 24; the value of the "Notify" error code is 25.

The following defines error values for the Routing Problem Error Code:

Value	Error:
1	Bad EXPLICIT_ROUTE object
2	Bad strict node
3	Bad loose node
4	Bad initial subobject
5	No route available toward destination
6	Unacceptable label value
7	RRO indicated routing loops
8	MPLS being negotiated, but a non-RSVP-capable router stands in the path
9	MPLS label allocation failure
10	Unsupported L3PID

For the Notify Error Code, the 16 bits of the Error Value field are:

ss00 cccc cccc cccc

The high order bits are as defined under Error Code 1. (See [1]).

When ss = 00, the following subcodes are defined:

- 1 RRO too large for MTU
- 2 RRO notification
- 3 Tunnel locally repaired

#### 4.6. Session, Sender Template, and Filter Spec Objects

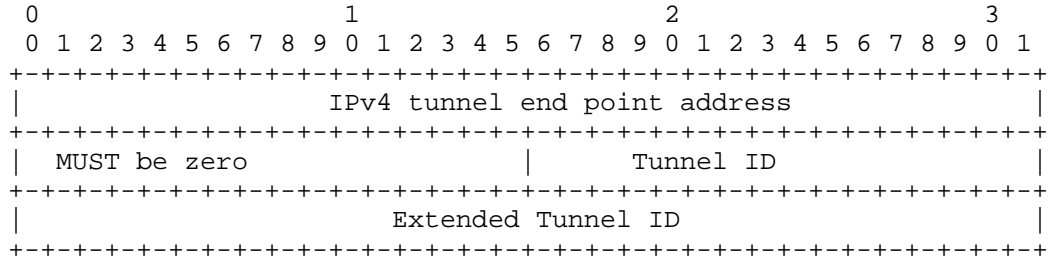
New C-Types are defined for the SESSION, SENDER\_TEMPLATE and FILTER\_SPEC objects.

The LSP\_TUNNEL objects have the following format:

4.6.1. Session Object

4.6.1.1. LSP\_TUNNEL\_IPv4 Session Object

Class = SESSION, LSP\_TUNNEL\_IPv4 C-Type = 7



IPv4 tunnel end point address

IPv4 address of the egress node for the tunnel.

Tunnel ID

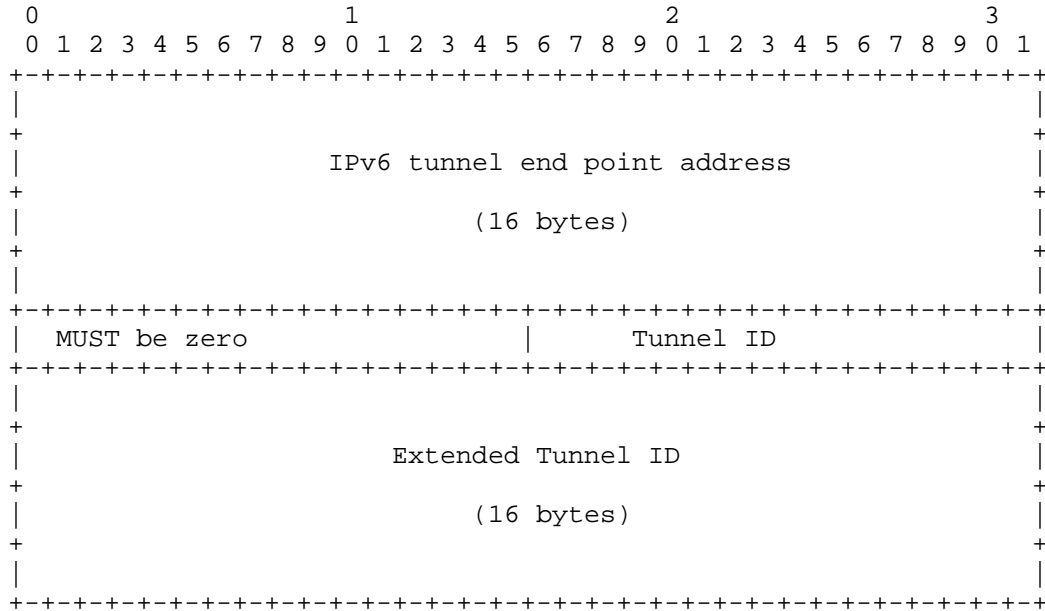
A 16-bit identifier used in the SESSION that remains constant over the life of the tunnel.

Extended Tunnel ID

A 32-bit identifier used in the SESSION that remains constant over the life of the tunnel. Normally set to all zeros. Ingress nodes that wish to narrow the scope of a SESSION to the ingress-egress pair may place their IPv4 address here as a globally unique identifier.

4.6.1.2. LSP\_TUNNEL\_IPv6 Session Object

Class = SESSION, LSP\_TUNNEL\_IPv6 C\_Type = 8



IPv6 tunnel end point address

IPv6 address of the egress node for the tunnel.

Tunnel ID

A 16-bit identifier used in the SESSION that remains constant over the life of the tunnel.

Extended Tunnel ID

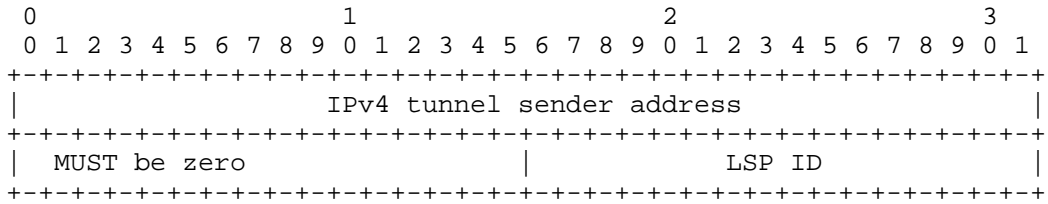
A 16-byte identifier used in the SESSION that remains constant over the life of the tunnel. Normally set to all zeros. Ingress nodes that wish to narrow the scope of a SESSION to the ingress-egress pair may place their IPv6 address here as a globally unique identifier.

4.6.2. Sender Template Object

4.6.2.1. LSP\_TUNNEL\_IPv4 Sender Template Object

Class = SENDER\_TEMPLATE, LSP\_TUNNEL\_IPv4 C-Type = 7





IPv4 tunnel sender address

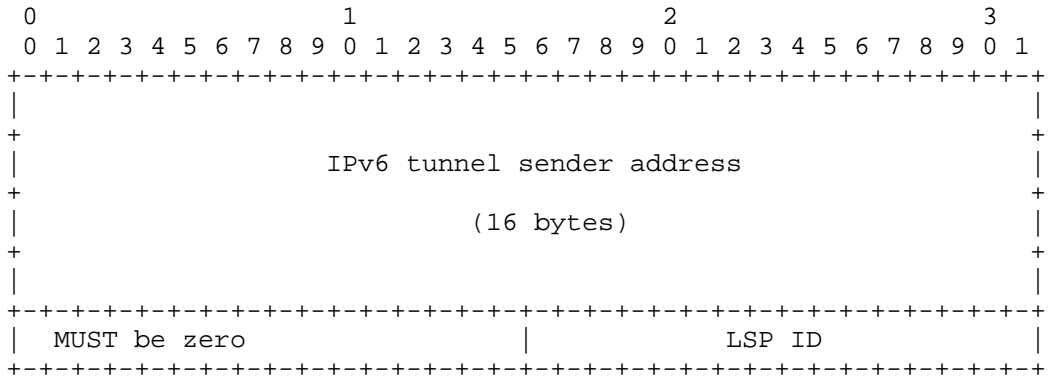
IPv4 address for a sender node

LSP ID

A 16-bit identifier used in the SENDER\_TEMPLATE and the FILTER\_SPEC that can be changed to allow a sender to share resources with itself.

4.6.2.2. LSP\_TUNNEL\_IPv6 Sender Template Object

Class = SENDER\_TEMPLATE, LSP\_TUNNEL\_IPv6 C\_Type = 8



IPv6 tunnel sender address

IPv6 address for a sender node

LSP ID

A 16-bit identifier used in the SENDER\_TEMPLATE and the FILTER\_SPEC that can be changed to allow a sender to share resources with itself.

#### 4.6.3. Filter Specification Object

##### 4.6.3.1. LSP\_TUNNEL\_IPv4 Filter Specification Object

Class = FILTER SPECIFICATION, LSP\_TUNNEL\_IPv4 C-Type = 7

The format of the LSP\_TUNNEL\_IPv4 FILTER\_SPEC object is identical to the LSP\_TUNNEL\_IPv4 SENDER\_TEMPLATE object.

##### 4.6.3.2. LSP\_TUNNEL\_IPv6 Filter Specification Object

Class = FILTER SPECIFICATION, LSP\_TUNNEL\_IPv6 C\_Type = 8

The format of the LSP\_TUNNEL\_IPv6 FILTER\_SPEC object is identical to the LSP\_TUNNEL\_IPv6 SENDER\_TEMPLATE object.

#### 4.6.4. Reroute and Bandwidth Increase Procedure

This section describes how to setup a tunnel that is capable of maintaining resource reservations (without double counting) while it is being rerouted or while it is attempting to increase its bandwidth. In the initial Path message, the ingress node forms a SESSION object, assigns a Tunnel\_ID, and places its IPv4 address in the Extended\_Tunnel\_ID. It also forms a SENDER\_TEMPLATE and assigns a LSP\_ID. Tunnel setup then proceeds according to the normal procedure.

On receipt of the Path message, the egress node sends a Resv message with the STYLE Shared Explicit toward the ingress node.

When an ingress node with an established path wants to change that path, it forms a new Path message as follows. The existing SESSION object is used. In particular the Tunnel\_ID and Extended\_Tunnel\_ID are unchanged. The ingress node picks a new LSP\_ID to form a new SENDER\_TEMPLATE. It creates an EXPLICIT\_ROUTE object for the new route. The new Path message is sent. The ingress node refreshes both the old and new path messages.

The egress node responds with a Resv message with an SE flow descriptor formatted as:

```
<FLOWSPEC><old_FILTER_SPEC><old_LABEL_OBJECT><new_FILTER_SPEC>  
<new_LABEL_OBJECT>
```

(Note that if the PHOPs are different, then two messages are sent each with the appropriate FILTER\_SPEC and LABEL\_OBJECT.)

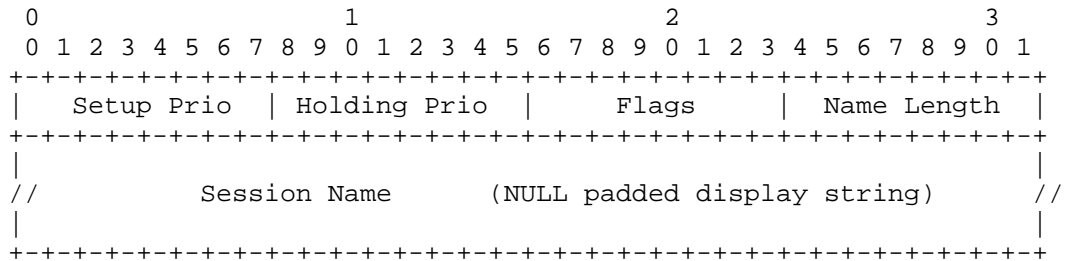
When the ingress node receives the Resv Message(s), it may begin using the new route. It SHOULD send a PathTear message for the old route.

4.7. Session Attribute Object

The Session Attribute Class is 207. Two C\_Types are defined, LSP\_TUNNEL, C-Type = 7 and LSP\_TUNNEL\_RA, C-Type = 1. The LSP\_TUNNEL\_RA C-Type includes all the same fields as the LSP\_TUNNEL C-Type. Additionally it carries resource affinity information. The formats are as follows:

4.7.1. Format without resource affinities

SESSION\_ATTRIBUTE class = 207, LSP\_TUNNEL C-Type = 7



Setup Priority

The priority of the session with respect to taking resources, in the range of 0 to 7. The value 0 is the highest priority. The Setup Priority is used in deciding whether this session can preempt another session.

Holding Priority

The priority of the session with respect to holding resources, in the range of 0 to 7. The value 0 is the highest priority. Holding Priority is used in deciding whether this session can be preempted by another session.

## Flags

0x01 Local protection desired

This flag permits transit routers to use a local repair mechanism which may result in violation of the explicit route object. When a fault is detected on an adjacent downstream link or node, a transit router can reroute traffic for fast service restoration.

0x02 Label recording desired

This flag indicates that label information should be included when doing a route record.

0x04 SE Style desired

This flag indicates that the tunnel ingress node may choose to reroute this tunnel without tearing it down. A tunnel egress node SHOULD use the SE Style when responding with a Resv message.

## Name Length

The length of the display string before padding, in bytes.

## Session Name

A null padded string of characters.

4.7.2. Format with resource affinities

```

SESSION_ATTRIBUTE class = 207, LSP_TUNNEL_RA C-Type = 1

      0                1                2                3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Exclude-any                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Include-any                                    |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Include-all                                 |
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Setup Prio | Holding Prio |      Flags      |  Name Length  |
+-----+-----+-----+-----+-----+-----+-----+-----+
| //          Session Name      (NULL padded display string)          //
|
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Exclude-any

A 32-bit vector representing a set of attribute filters associated with a tunnel any of which renders a link unacceptable.

Include-any

A 32-bit vector representing a set of attribute filters associated with a tunnel any of which renders a link acceptable (with respect to this test). A null set (all bits set to zero) automatically passes.

Include-all

A 32-bit vector representing a set of attribute filters associated with a tunnel all of which must be present for a link to be acceptable (with respect to this test). A null set (all bits set to zero) automatically passes.

Setup Priority

The priority of the session with respect to taking resources, in the range of 0 to 7. The value 0 is the highest priority. The Setup Priority is used in deciding whether this session can preempt another session.

#### Holding Priority

The priority of the session with respect to holding resources, in the range of 0 to 7. The value 0 is the highest priority. Holding Priority is used in deciding whether this session can be preempted by another session.

#### Flags

0x01 Local protection desired

This flag permits transit routers to use a local repair mechanism which may result in violation of the explicit route object. When a fault is detected on an adjacent downstream link or node, a transit router can reroute traffic for fast service restoration.

0x02 Label recording desired

This flag indicates that label information should be included when doing a route record.

0x04 SE Style desired

This flag indicates that the tunnel ingress node may choose to reroute this tunnel without tearing it down. A tunnel egress node SHOULD use the SE Style when responding with a Resv message.

#### Name Length

The length of the display string before padding, in bytes.

#### Session Name

A null padded string of characters.

#### 4.7.3. Procedures applying to both C-Types

The support of setup and holding priorities is OPTIONAL. A node can recognize this information but be unable to perform the requested operation. The node SHOULD pass the information downstream unchanged.

As noted above, preemption is implemented by two priorities. The Setup Priority is the priority for taking resources. The Holding Priority is the priority for holding a resource. Specifically, the

Holding Priority is the priority at which resources assigned to this session will be reserved. The Setup Priority SHOULD never be higher than the Holding Priority for a given session.

The setup and holding priorities are directly analogous to the preemption and defending priorities as defined in [9]. While the interaction of these two objects is ultimately a matter of policy, the following default interaction is RECOMMENDED.

When both objects are present, the preemption priority policy element is used. A mapping between the priority spaces is defined as follows. A session attribute priority S is mapped to a preemption priority P by the formula  $P = 2^{(14-2S)}$ . The reverse mapping is shown in the following table.

Preemption Priority	Session Attribute Priority
0 - 3	7
4 - 15	6
16 - 63	5
64 - 255	4
256 - 1023	3
1024 - 4095	2
4096 - 16383	1
16384 - 65535	0

When a new Path message is considered for admission, the bandwidth requested is compared with the bandwidth available at the priority specified in the Setup Priority.

If the requested bandwidth is not available a PathErr message is returned with an Error Code of 01, Admission Control Failure, and an Error Value of 0x0002. The first 0 in the Error Value indicates a globally defined subcode and is not informational. The 002 indicates "requested bandwidth unavailable".

If the requested bandwidth is less than the unused bandwidth then processing is complete. If the requested bandwidth is available, but is in use by lower priority sessions, then lower priority sessions (beginning with the lowest priority) MAY be preempted to free the necessary bandwidth.

When preemption is supported, each preempted reservation triggers a TC\_Preempt() upcall to local clients, passing a subcode that indicates the reason. A ResvErr and/or PathErr with the code "Policy Control failure" SHOULD be sent toward the downstream receivers and upstream senders.

The support of local-protection is OPTIONAL. A node may recognize the local-protection Flag but may be unable to perform the requested operation. In this case, the node SHOULD pass the information downstream unchanged.

The recording of the Label subobject in the ROUTE\_RECORD object is controlled by the label-recording-desired flag in the SESSION\_ATTRIBUTE object. Since the Label subobject is not needed for all applications, it is not automatically recorded. The flag allows applications to request this only when needed.

The contents of the Session Name field are a string, typically of display-able characters. The Length MUST always be a multiple of 4 and MUST be at least 8. For an object length that is not a multiple of 4, the object is padded with trailing NULL characters. The Name Length field contains the actual string length.

#### 4.7.4. Resource Affinity Procedures

Resource classes and resource class affinities are described in [3]. In this document we use the briefer term resource affinities for the latter term. Resource classes can be associated with links and advertised in routing protocols. Resource class affinities are used by RSVP in two ways. In order to be validated a link MUST pass the three tests below. If the test fails a PathErr with the code "policy control failure" SHOULD be sent.

When a new reservation is considered for admission over a strict node in an ERO, a node MAY validate the resource affinities with the resource classes of that link. When a node is choosing links in order to extend a loose node of an ERO, the node MUST validate the resource classes of those links against the resource affinities. If no acceptable links can be found to extend the ERO, the node SHOULD send a PathErr message with an error code of "Routing Problem" and an error value of "no route available toward destination".

In order to be validated a link MUST pass the following three tests.

To precisely describe the tests use the definitions in the object description above. We also define

Link-attr            A 32-bit vector representing attributes associated with a link.



The three tests are

1. Exclude-any

This test excludes a link from consideration if the link carries any of the attributes in the set.

```
(link-attr & exclude-any) == 0
```

2. Include-any

This test accepts a link if the link carries any of the attributes in the set.

```
(include-any == 0) | ((link-attr & include-any) != 0)
```

3. Include-all

This test accepts a link only if the link carries all of the attributes in the set.

```
(include-all == 0) | (((link-attr & include-all) ^ include-all) == 0)
```

For a link to be acceptable, all three tests MUST pass. If the test fails, the node SHOULD send a PathErr message with an error code of "Routing Problem" and an error value of "no route available toward destination".

If a Path message contains multiple SESSION\_ATTRIBUTE objects, only the first SESSION\_ATTRIBUTE object is meaningful. Subsequent SESSION\_ATTRIBUTE objects can be ignored and need not be forwarded.

All RSVP routers, whether they support the SESSION\_ATTRIBUTE object or not, SHALL forward the object unmodified. The presence of non-RSVP routers anywhere between senders and receivers has no impact on this object.

5. Hello Extension

The RSVP Hello extension enables RSVP nodes to detect when a neighboring node is not reachable. The mechanism provides node to node failure detection. When such a failure is detected it is handled much the same as a link layer communication failure. This mechanism is intended to be used when notification of link layer failures is not available and unnumbered links are not used, or when the failure detection mechanisms provided by the link layer are not sufficient for timely node failure detection.

It should be noted that node failure detection is not the same as a link failure detection mechanism, particularly in the case of multiple parallel unnumbered links.

The Hello extension is specifically designed so that one side can use the mechanism while the other side does not. Neighbor failure detection may be initiated at any time. This includes when neighbors first learn about each other, or just when neighbors are sharing Resv or Path state.

The Hello extension is composed of a Hello message, a HELLO REQUEST object and a HELLO ACK object. Hello processing between two neighbors supports independent selection of, typically configured, failure detection intervals. Each neighbor can autonomously issue HELLO REQUEST objects. Each request is answered by an acknowledgment. Hello Messages also contain enough information so that one neighbor can suppress issuing hello requests and still perform neighbor failure detection. A Hello message may be included as a sub-message within a bundle message.

Neighbor failure detection is accomplished by collecting and storing a neighbor's "instance" value. If a change in value is seen or if the neighbor is not properly reporting the locally advertised value, then the neighbor is presumed to have reset. When a neighbor's value is seen to change or when communication is lost with a neighbor, then the instance value advertised to that neighbor is also changed. The HELLO objects provide a mechanism for polling for and providing an instance value. A poll request also includes the sender's instance value. This allows the receiver of a poll to optionally treat the poll as an implicit poll response. This optional handling is an optimization that can reduce the total number of polls and responses processed by a pair of neighbors. In all cases, when both sides support the optimization the result will be only one set of polls and responses per failure detection interval. Depending on selected intervals, the same benefit can occur even when only one neighbor supports the optimization.

#### 5.1. Hello Message Format

Hello Messages are always sent between two RSVP neighbors. The IP source address is the IP address of the sending node. The IP destination address is the IP address of the neighbor node.

The HELLO mechanism is intended for use between immediate neighbors. When HELLO messages are being exchanged between immediate neighbors, the IP TTL field of all outgoing HELLO messages SHOULD be set to 1.

The Hello message has a Msg Type of 20. The Hello message format is as follows:

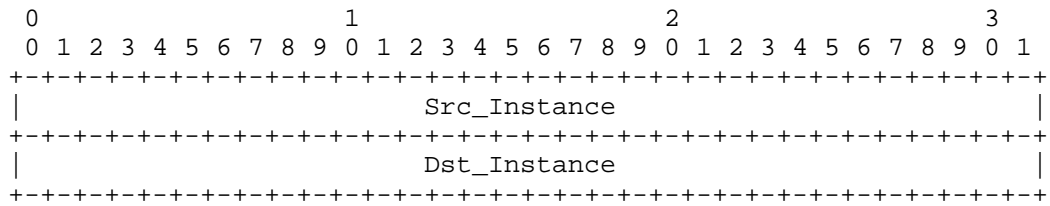
```
<Hello Message> ::= <Common Header> [ <INTEGRITY> ]
                    <HELLO>
```

5.2. HELLO Object formats

The HELLO Class is 22. There are two C\_Types defined.

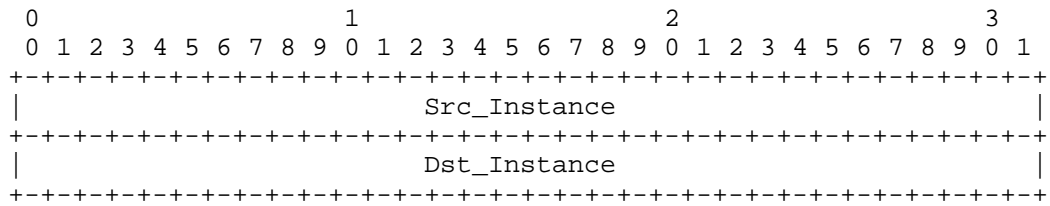
5.2.1. HELLO REQUEST object

Class = HELLO Class, C\_Type = 1



5.2.2. HELLO ACK object

Class = HELLO Class, C\_Type = 2



Src\_Instance: 32 bits

a 32 bit value that represents the sender's instance. The advertiser maintains a per neighbor representation/value. This value MUST change when the sender is reset, when the node reboots, or when communication is lost to the neighboring node and otherwise remains the same. This field MUST NOT be set to zero (0).

Dst\_Instance: 32 bits

The most recently received Src\_Instance value received from the neighbor. This field MUST be set to zero (0) when no value has ever been seen from the neighbor.

### 5.3. Hello Message Usage

The Hello Message is completely OPTIONAL. All messages may be ignored by nodes which do not wish to participate in Hello message processing. The balance of this section is written assuming that the receiver as well as the sender is participating. In particular, the use of MUST and SHOULD with respect to the receiver applies only to a node that supports Hello message processing.

A node periodically generates a Hello message containing a HELLO REQUEST object for each neighbor who's status is being tracked. The periodicity is governed by the `hello_interval`. This value MAY be configured on a per neighbor basis. The default value is 5 ms.

When generating a message containing a HELLO REQUEST object, the sender fills in the `Src_Instance` field with a value representing it's per neighbor instance. This value MUST NOT change while the agent is exchanging Hellos with the corresponding neighbor. The sender also fills in the `Dst_Instance` field with the `Src_Instance` value most recently received from the neighbor. For reference, call this variable `Neighbor_Src_Instance`. If no value has ever been received from the neighbor or this node considers communication to the neighbor to have been lost, the `Neighbor_Src_Instance` is set to zero (0). The generation of a message SHOULD be suppressed when a HELLO REQUEST object was received from the destination node within the prior `hello_interval` interval.

On receipt of a message containing a HELLO REQUEST object, the receiver MUST generate a Hello message containing a HELLO ACK object. The receiver SHOULD also verify that the neighbor has not reset. This is done by comparing the sender's `Src_Instance` field value with the previously received value. If the `Neighbor_Src_Instance` value is zero, and the `Src_Instance` field is non-zero, the `Neighbor_Src_Instance` is updated with the new value. If the value differs or the `Src_Instance` field is zero, then the node MUST treat the neighbor as if communication has been lost.

The receiver of a HELLO REQUEST object SHOULD also verify that the neighbor is reflecting back the receiver's Instance value. This is done by comparing the received `Dst_Instance` field with the `Src_Instance` field value most recently transmitted to that neighbor. If the neighbor continues to advertise a wrong non-zero value after a configured number of intervals, then the node MUST treat the neighbor as if communication has been lost.

On receipt of a message containing a HELLO ACK object, the receiver MUST verify that the neighbor has not reset. This is done by comparing the sender's `Src_Instance` field value with the previously

received value. If the Neighbor\_Src\_Instance value is zero, and the Src\_Instance field is non-zero, the Neighbor\_Src\_Instance is updated with the new value. If the value differs or the Src\_Instance field is zero, then the node MUST treat the neighbor as if communication has been lost.

The receiver of a HELLO ACK object MUST also verify that the neighbor is reflecting back the receiver's Instance value. If the neighbor advertises a wrong value in the Dst\_Instance field, then a node MUST treat the neighbor as if communication has been lost.

If no Instance values are received, via either REQUEST or ACK objects, from a neighbor within a configured number of hello\_intervals, then a node MUST presume that it cannot communicate with the neighbor. The default for this number is 3.5.

When communication is lost or presumed to be lost as described above, a node MAY re-initiate HELLOs. If a node does re-initiate it MUST use a Src\_Instance value different than the one advertised in the previous HELLO message. This new value MUST continue to be advertised to the corresponding neighbor until a reset or reboot occurs, or until another communication failure is detected. If a new instance value has not been received from the neighbor, then the node MUST advertise zero in the Dst\_instance value field.

#### 5.4. Multi-Link Considerations

As previously noted, the Hello extension is targeted at detecting node failures not per link failures. When there is only one link between neighboring nodes or when all links between a pair of nodes fail, the distinction between node and link failures is not really meaningful and handling of such failures has already been covered. When there are multiple links shared between neighbors, there are special considerations. When the links between neighbors are numbered, then Hellos MUST be run on each link and the previously described mechanisms apply.

When the links are unnumbered, link failure detection MUST be provided by some means other than Hellos. Each node SHOULD use a single Hello exchange with the neighbor. The case where all links have failed, is the same as the no received value case mentioned in the previous section.

### 5.5. Compatibility

The Hello extension does not affect the processing of any other RSVP message. The only effect is to allow a link (node) down event to be declared sooner than it would have been. RSVP response to that condition is unchanged.

The Hello extension is fully backwards compatible. The Hello class is assigned a class value of the form 0bbbbbbb. Depending on the implementation, implementations that do not support the extension will either silently discard Hello messages or will respond with an "Unknown Object Class" error. In either case the sender will fail to see an acknowledgment for the issued Hello.

### 6. Security Considerations

In principle these extensions to RSVP pose no security exposures over and above RFC 2205[1]. However, there is a slight change in the trust model. Traffic sent on a normal RSVP session can be filtered according to source and destination addresses as well as port numbers. In this specification, filtering occurs only on the basis of an incoming label. For this reason an administration may wish to limit the domain over which LSP tunnels can be established. This can be accomplished by setting filters on various ports to deny action on a RSVP path message with a SESSION object of type LSP\_TUNNEL\_IPv4 (7) or LSP\_TUNNEL\_IPv6 (8).

### 7. IANA Considerations

IANA assigns values to RSVP protocol parameters. Within the current document an EXPLICIT\_ROUTE object and a ROUTE\_RECORD object are defined. Each of these objects contain subobjects. This section defines the rules for the assignment of subobject numbers. This section uses the terminology of BCP 26 "Guidelines for Writing an IANA Considerations Section in RFCs" [15].

#### EXPLICIT\_ROUTE Subobject Type

EXPLICIT\_ROUTE Subobject Type is a 7-bit number that identifies the function of the subobject. There are no range restrictions. All possible values are available for assignment.

Following the policies outlined in [15], subobject types in the range 0 - 63 (0x00 - 0x3F) are allocated through an IETF Consensus action, codes in the range 64 - 95 (0x40 - 0x5F) are allocated as First Come First Served, and codes in the range 96 - 127 (0x60 - 0x7F) are reserved for Private Use.

## ROUTE\_RECORD Subobject Type

ROUTE\_RECORD Subobject Type is an 8-bit number that identifies the function of the subobject. There are no range restrictions. All possible values are available for assignment.

Following the policies outlined in [15], subobject types in the range 0 - 127 (0x00 - 0x7F) are allocated through an IETF Consensus action, codes in the range 128 - 191 (0x80 - 0xBF) are allocated as First Come First Served, and codes in the range 192 - 255 (0xC0 - 0xFF) are reserved for Private Use.

The following assignments are made in this document.

## 7.1. Message Types

Message Number	Message Name
----------------	--------------

20	Hello
----	-------

## 7.2. Class Numbers and C-Types

Class Number	Class Name
--------------	------------

1	SESSION
---	---------

Class Types or C-Types:

7	LSP Tunnel IPv4
8	LSP Tunnel IPv6

10	FILTER_SPEC
----	-------------

Class Types or C-Types:

7	LSP Tunnel IPv4
8	LSP Tunnel IPv6

11	SENDER_TEMPLATE
----	-----------------

Class Types or C-Types:

7	LSP Tunnel IPv4
8	LSP Tunnel IPv6

- 16    RSVP\_LABEL
- Class Types or C-Types:
- 1        Type 1 Label
- 19    LABEL\_REQUEST
- Class Types or C-Types:
- 1        Without Label Range
- 2        With ATM Label Range
- 3        With Frame Relay Label Range
- 20    EXPLICIT\_ROUTE
- Class Types or C-Types:
- 1        Type 1 Explicit Route
- 21    ROUTE\_RECORD
- Class Types or C-Types:
- 1        Type 1 Route Record
- 22    HELLO
- Class Types or C-Types:
- 1        Request
- 2        Acknowledgment
- 207   SESSION\_ATTRIBUTE
- Class Types or C-Types:
- 1        LSP\_TUNNEL\_RA
- 7        LSP Tunnel



### 7.3. Error Codes and Globally-Defined Error Value Sub-Codes

The following list extends the basic list of Error Codes and Values that are defined in [RFC2205].

Error Code	Meaning
------------	---------

24	Routing Problem
----	-----------------

This Error Code has the following globally-defined Error Value sub-codes:

1	Bad EXPLICIT_ROUTE object
2	Bad strict node
3	Bad loose node
4	Bad initial subobject
5	No route available toward destination
6	Unacceptable label value
7	RRO indicated routing loops
8	MPLS being negotiated, but a non-RSVP-capable router stands in the path
9	MPLS label allocation failure
10	Unsupported L3PID

25	Notify Error
----	--------------

This Error Code has the following globally-defined Error Value sub-codes:

1	RRO too large for MTU
2	RRO Notification
3	Tunnel locally repaired

### 7.4. Subobject Definitions

Subobjects of the EXPLICIT\_ROUTE object with C-Type 1:

1	IPv4 prefix
2	IPv6 prefix
32	Autonomous system number

Subobjects of the RECORD\_ROUTE object with C-Type 1:

- 1 IPv4 address
- 2 IPv6 address
- 3 Label

## 8. Intellectual Property Considerations

The IETF has been notified of intellectual property rights claimed in regard to some or all of the specification contained in this document. For more information consult the online list of claimed rights.

## 9. Acknowledgments

This document contains ideas as well as text that have appeared in previous Internet Drafts. The authors of the current document wish to thank the authors of those drafts. They are Steven Blake, Bruce Davie, Roch Guerin, Sanjay Kamat, Yakov Rekhter, Eric Rosen, and Arun Viswanathan. We also wish to thank Bora Akyol, Yoram Bernet and Alex Mondrus for their comments on this document.

## 10. References

- [1] Braden, R., Zhang, L., Berson, S., Herzog, S. and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1, Functional Specification", RFC 2205, September 1997.
- [2] Rosen, E., Viswanathan, A. and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.
- [3] Awduche, D., Malcolm, J., Agogbua, J., O'Dell and J. McManus, "Requirements for Traffic Engineering over MPLS", RFC 2702, September 1999.
- [4] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, September 1997.
- [5] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T. and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.
- [6] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [7] Almquist, P., "Type of Service in the Internet Protocol Suite", RFC 1349, July 1992.

- [8] Nichols, K., Blake, S., Baker, F. and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.
- [9] Herzog, S., "Signaled Preemption Priority Policy Element", RFC 2751, January 2000.
- [10] Awduche, D., Hannan, A. and X. Xiao, "Applicability Statement for Extensions to RSVP for LSP-Tunnels", RFC 3210, December 2001.
- [11] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", RFC 2210, September 1997.
- [12] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, September 1981.
- [13] Mogul, J. and S. Deering, "Path MTU Discovery", RFC 1191, November 1990.
- [14] Conta, A. and S. Deering, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6)", RFC 2463, December 1998.
- [15] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.
- [16] Bernet, Y., Smit, A. and B. Davie, "Specification of the Null Service Type", RFC 2997, November 2000.

## 11. Authors' Addresses

Daniel O. Awduche  
Movaz Networks, Inc.  
7926 Jones Branch Drive, Suite 615  
McLean, VA 22102  
Voice: +1 703-298-5291  
EMail: awduche@movaz.com

Lou Berger  
Movaz Networks, Inc.  
7926 Jones Branch Drive, Suite 615  
McLean, VA 22102  
Voice: +1 703 847 1801  
EMail: lberger@movaz.com

Der-Hwa Gan  
Juniper Networks, Inc.  
385 Ravendale Drive  
Mountain View, CA 94043  
EMail: dhg@juniper.net

Tony Li  
Procket Networks  
3910 Freedom Circle, Ste. 102A  
Santa Clara CA 95054  
EMail: tli@procket.com

Vijay Srinivasan  
Cosine Communications, Inc.  
1200 Bridge Parkway  
Redwood City, CA 94065  
Voice: +1 650 628 4892  
EMail: vsriniva@cosinecom.com

George Swallow  
Cisco Systems, Inc.  
250 Apollo Drive  
Chelmsford, MA 01824  
Voice: +1 978 244 8143  
EMail: swallow@cisco.com

## 12. Full Copyright Statement

Copyright (C) The Internet Society (2001). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

## Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.

