

White Paper

Multiprotocol Label Switching

Enhancing Routing in the New Public Network

Chuck Semeria
Marketing Engineer



Juniper Networks, Inc.
385 Ravendale Drive
Mountain View, CA 94043 USA
650-526-8000
www.juniper.net

Part Number : 200001-002 9/27/99

Contents

Executive Summary	3
Perspective	3
The Evolution of Multilayer Switching in the Internet	4
Fundamental Building Blocks	4
Separation of Control and Forwarding Components	4
Label-Swapping Forwarding Algorithm	5
ISPs Migrate to the IP-over-ATM Model	7
Multilayer Switching Alternatives to the IP-over-ATM Model	8
Similarities Among the Multilayer Switching Solutions	8
Fundamental Differences Among the Multilayer Switching Solutions	10
Data-Driven Model	10
Control-Driven Model	11
Fundamental Problem with Multilayer Switching Solutions	11
Multiprotocol Label Switching	12
Requirements and Objectives	13
Common Misconceptions about MPLS	13
The Promise of MPLS	14
MPLS Applications	16
Traffic Engineering	16
Class of Service	17
Virtual Private Networks	17
Conclusion	18
Related Reading	19
References	19
Textbooks	19
Requests for Comments	19
Internet Drafts	19
Next Generation Networks 1998: Conference Tutorials and Sessions	20
UUNET IW-MPLS '98 Conference	20

List of Figures

Figure 1: Routing Functional Components: Control and Forwarding	5
Figure 2: Packet Traversing a Label Switched Path	6
Figure 3: IP-over-ATM Model	7
Figure 4: Multilayer Switch as a Fast IP Router	8
Figure 5: Multilayer Switching Solutions—Integrating IP Routing and ATM Switching	10
Figure 6: Multiprotocol Label Switching	12
Figure 7: MPLS Header	13
Figure 8: MPLS Enhances Routing Functionality	14
Figure 9: How MPLS Enhances Routing Functionality	15
Figure 10: Traffic Engineering LSP vs. IGP Shortest Path across a Service Provider's Network	16
Figure 11: MPLS Facilitates the Deployment of VPNs	18

Executive Summary

Both large and small Internet Service Providers (ISPs) constantly face the challenges of adapting their networks to support rapid growth and customer demand for more reliable and differentiated services. In the mid-1990s, the IP-over-ATM model provided many ISPs a solution for delivering excellent performance and performing traffic engineering. Moreover, many carriers found it cost-effective to multiplex Internet traffic as one of many services carried over an ATM core.

Recently, the growth of Internet services and Wavelength Division Multiplexing (WDM) technology at the fiber level have provided a viable alternative to ATM for multiplexing multiple services over individual circuits. In addition, the once faster and higher bandwidth ATM switches are being out-performed by Internet backbone routers. Equally important, Multiprotocol Label Switching (MPLS) offers simpler mechanisms for packet-oriented traffic engineering and multiservice functionality with the added benefit of greater scalability.

MPLS emerged from the IETF's effort to standardize a number of proprietary multilayer switching solutions that were initially proposed in the mid-1990s. To help you appreciate the importance of MPLS and its impact on the Internet core, the first half of this paper describes the forces that motivated the development and evolution of these different solutions, focusing on the common features and design considerations shared by the different solutions—the complete separation of the control component from the forwarding component and the use of a label-swapping forwarding paradigm. This section also describes the natural technological evolution that took place and that eventually culminated in the IETF working group's definition of MPLS.

The second half of this paper builds on your understanding of multilayer switching, focusing on the specifics of MPLS. It describes the goals and objectives of the MPLS working group, the core MPLS components, some of the common misconceptions about MPLS, the benefits of MPLS for the core of the Internet, and the most popular applications for MPLS. This section describes how MPLS is the foundation for service differentiation because it permits ISPs to deliver new services that cannot be readily supported by conventional IP routing techniques.

Perspective

Over the past few years, a number of new technologies have been designed to support Internet Service Providers (ISPs) as they try to keep a step ahead of the Internet's explosive growth. The latest technological advances include Internet backbone routers, new queuing and scheduling algorithms, IPSEC, web-caching services, directory services, and integrated routing/forwarding solutions. While all these technologies are critical for the successful operation and continued growth of the Internet, the evolution of routing functionality is essential if ISPs want to provide support for a new class of revenue-generating customer services.

Multiprotocol Label Switching (MPLS) is the latest step in the evolution of routing/forwarding technology for the core of the Internet. MPLS delivers a solution that seamlessly integrates the control of IP routing with the simplicity of Layer 2 switching. Furthermore, MPLS provides a foundation that supports the deployment of advanced routing services because it solves a number of complex problems:

- ▶ MPLS addresses the scalability issues associated with the currently deployed IP-over-ATM overlay model.
- ▶ MPLS significantly reduces the complexity of network operation.
- ▶ MPLS facilitates the delivery of new routing capabilities that enhance conventional IP routing techniques.
- ▶ MPLS offers a standards-based solution that promotes multivendor interoperability.

MPLS emerged from the IETF's effort to standardize a set of proprietary multilayer switching solutions that were originally developed in the mid-1990s. To fully understand the essence of MPLS and its role in the Internet, it is valuable to look back and examine the forces that stimulated the development of these proprietary multilayer switching approaches and how they were ultimately integrated into MPLS.

The Evolution of Multilayer Switching in the Internet

Multilayer switching describes the integration of Layer 2 switching and Layer 3 routing. Today, some ISP networks are built using an *overlay model* in which a logical IP routed topology runs over and is independent of an underlying Layer 2 switched topology (ATM or Frame Relay). Layer 2 switches provide high-speed connectivity, while the IP routers at the edge—interconnected by a mesh of Layer 2 virtual circuits—provide the intelligence to forward IP datagrams. The difficulty with this approach lies in the complexity of mapping between two distinct architectures that require the definition and maintenance of separate topologies, address spaces, routing protocols, signaling protocols, and resource allocation schemes. The emergence of the multilayer switching solutions and MPLS is part of the evolution of the Internet to decrease complexity by combining Layer 2 switching and Layer 3 routing into a fully integrated solution.

Fundamental Building Blocks

Before beginning our discussion of multilayer switching in the Internet, it is important to understand the fundamental building blocks common to all multilayer switching solutions and MPLS:

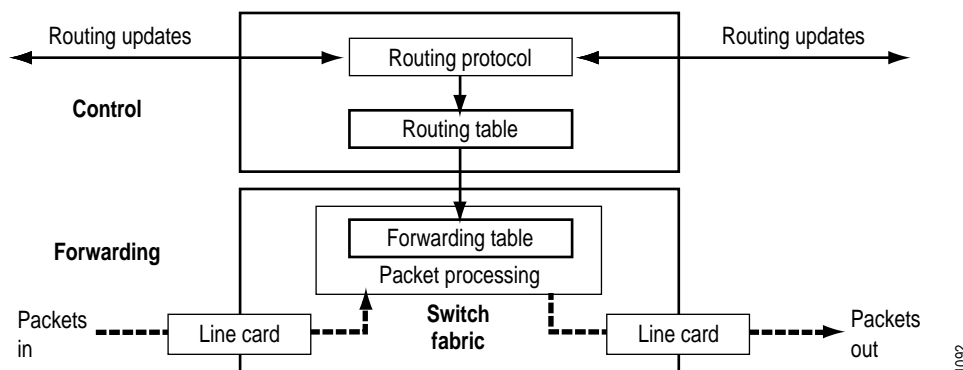
- ▶ Separation of the control and forwarding components.
- ▶ Label-swapping forwarding algorithm.

Separation of Control and Forwarding Components

All multilayer switching solutions, including MPLS, are composed of two distinct functional components—a control component and a forwarding component (see Figure 1). The control component uses standard routing protocols (OSPF, IS-IS, and BGP-4) to exchange information with other routers to build and maintain a forwarding table. When packets arrive, the forwarding component searches the forwarding table maintained by the control component to make a routing decision for each packet. Specifically, the forwarding component examines

information contained in the packet's header, searches the forwarding table for a match, and directs the packet from the input interface to the output interface across the system's switching fabric.

Figure 1: Routing Functional Components: Control and Forwarding



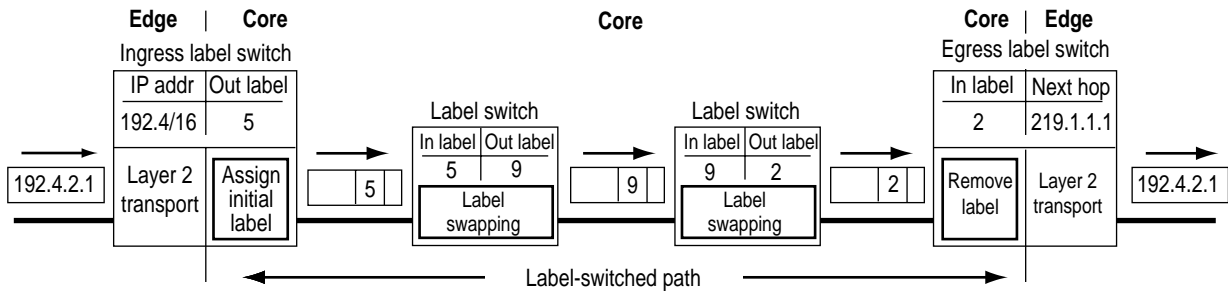
By completely separating the control component from the forwarding component, each component can be independently developed and modified. The only requirement is that the control component continue to communicate with the forwarding component by managing the packet-forwarding table. We will see that the deployment of an extremely simple forwarding algorithm, such as label swapping, can provide the extended forwarding capabilities needed to support new revenue-generating customer services.

Label-Swapping Forwarding Algorithm

The forwarding component of virtually all multilayer switching solutions and MPLS is based on a label-swapping forwarding algorithm. This is the same algorithm used to forward data in ATM and Frame Relay switches. Signaling and label distribution are fundamental to the operation of the label-swapping forwarding algorithm, but they are not discussed in this white paper. However, they will be discussed in a future Juniper Networks white paper.

A label is a short, fixed-length value carried in the packet's header to identify a Forwarding Equivalence Class (FEC). A label is analogous to a connection identifier, such as an ATM VPI/VCI or a Frame Relay DLCI, because it has only link-local significance, does not encode information from the network layer header, and maps traffic to a specific FEC. An FEC is a set of packets that are forwarded over the same path through a network even if their ultimate destinations are different. For example, in conventional longest-match IP routing, the set of unicast packets whose destination addresses map to a given IP address prefix is an example of an FEC.

The label-swapping forwarding algorithm requires packet classification at the ingress edge of the network to assign an initial label to each packet. In Figure 2, the ingress label switch receives an unlabeled packet with a destination address of 192.4.2.1. The label switch performs a longest-match routing table lookup and maps the packet to an FEC—192.4/16. The ingress label switch then assigns a label (with a value of 5) to the packet and forwards it to the next hop in the label-switched path (LSP).

Figure 2: Packet Traversing a Label Switched Path


1093

An LSP is functionally equivalent to a virtual circuit because it defines an ingress-to-egress path through a network that is followed by all packets assigned to a specific FEC. The first label switch in an LSP is called the ingress, or head-end, label switch. The last label switch in an LSP is called the egress, or tail-end, label switch.

In the core of the network, label switches ignore the packet's network layer header and simply forward the packet using the label-swapping algorithm. When a labeled packet arrives at a switch, the forwarding component uses the input port number and label to perform an exact match search of its forwarding table. When a match is found, the forwarding component retrieves the outgoing label, the outgoing interface, and the next-hop address from the forwarding table. The forwarding component then swaps (or replaces) the incoming label with the outgoing label and directs the packet to the outbound interface for transmission to the next hop in the LSP.

When the labeled packet arrives at the egress label switch, the forwarding component searches its forwarding table. If the next hop is not a label switch, the egress switch discards the label and forwards the packet using conventional longest-match IP forwarding.

Label swapping provides a significant number of operational benefits when compared to conventional hop-by-hop network layer routing:

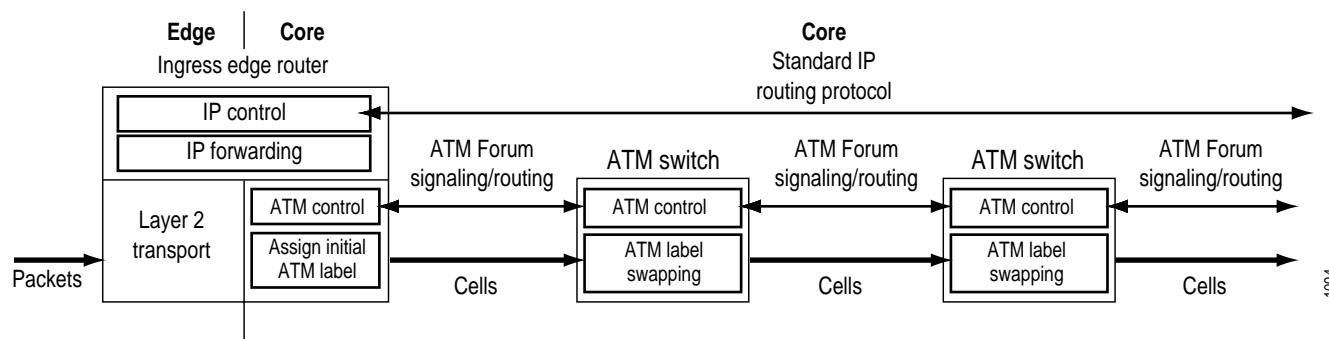
- Label swapping gives a service provider tremendous flexibility in the way that it assigns packets to FECs. For example, to simulate conventional IP forwarding, the ingress label switch can be configured to assign a packet to an FEC based on its destination address. However, packets can also be assigned to an FEC based on an unlimited number of policy-based considerations—the source address alone, the application type, the point of entry into the label-swapping network, the point of exit from the label-swapping network, the CoS conveyed in the packet header, or any combination of the above.
- Service providers can construct customized LSPs that support specific application requirements. LSPs can be designed to minimize the number of hops, meet certain bandwidth requirements, support precise performance requirements, bypass potential points of congestion, direct traffic away from the default path selected by the IGP, or simply force traffic across certain links or nodes in the network.
- The most important benefit of the label-swapping forwarding algorithm is its ability to take any type of user traffic, associate it with an FEC, and map the FEC to an LSP that has been specifically designed to satisfy the FEC's requirements. The deployment of technologies based on label-swapping forwarding techniques offer ISPs precise control over the flow of traffic in their networks. This unprecedented level of control results in a network that operates more efficiently and provides more predictable service.

ISPs Migrate to the IP-over-ATM Model

In the mid-1990s, certain ISPs evolved their networks from router-based cores to the overlay model of running IP over ATM. ISPs undertook this migration because they needed greater bandwidth, deterministic forwarding performance, and traffic engineering to support the explosive growth occurring in their networks. One of the primary reasons that the IP-over-ATM overlay model was able to satisfy these operational requirements was ATM’s use of a label-swapping forwarding algorithm.

The IP-over-ATM model (see Figure 3) was centered around ATM functionality, requiring ATM software controls (signaling and routing) and hardware forwarding (label swapping) on every system in the core of the network. The IP-over-ATM model met application requirements by using Layer 3 functionality at the edges of the network and maximized network throughput by relying on high-speed, label-swapping ATM switches and PVCs in the core. The role of IP routing was limited to the edges of the network because this model viewed software-based routers as the key source of poor network performance.

Figure 3: IP-over-ATM Model



In the mid-1990s, networking equipment was not specifically designed for Internet backbone applications so ISPs were forced to adapt whatever equipment was commercially available to support their rapidly growing networks. Only the ATM switching infrastructure equipment provided the bandwidth and forwarding capacity to support their immediate requirements. However, as ISPs continued their exponential growth and purpose-built equipment became available for Internet applications, continuing with the IP-over-ATM model and its inherent scalability problems made less sense. The scalability problems included the bandwidth limitations of ATM SAR interfaces, the 20 percent cell tax, the “n-squared” PVC problem, the Interior Gateway Protocol (IGP) stress, and the inability to operate over a non-ATM infrastructure.

Despite these scaling issues, the most challenging problem was the complexity of operating a network based on two disparate technologies that were independently designed and developed for entirely different tasks. IP and ATM are based on completely different protocol architectures (connectionless vs. connection-oriented), and they each have their own addressing models, routing protocols, signaling protocols, and resource allocation schemes. While rapidly growing ISPs required the performance and control that ATM and label-swapping provided, they realized that in a packet-based network environment, it became more difficult to justify the complexity of the IP-over-ATM model. Today, when high-performance Internet backbone routers are purpose-built for the core of the Internet, there are few good reasons to continue with an overly complex approach that requires two separate sets of equipment.

Multilayer Switching Alternatives to the IP-over-ATM Model

As ISPs continued migrating to the IP-over-ATM model, a number of technical, marketing, and financial trends began to influence the development of new technologies designed for the core of the Internet. The general public began to understand the Internet's prominent role in providing the foundation for a new global economy. It became clear that the Internet market was large enough to build equipment specifically designed for Internet backbone applications. IP quickly became the only protocol that mattered, winning out over IPX, AppleTalk, OSI, and SNA.

The notion of "IP convergence" provided venture capitalists and start-ups with a window of opportunity to compete with incumbent vendors. To be successful, a start-up needed to deliver a solution that provided the price and performance of an ATM switch and the control of an IP router, while eliminating the complex mapping required by the IP-over-ATM model. By late 1996, a number of vendors were promoting proprietary multilayer switching solutions that integrated ATM switching and IP routing, including:

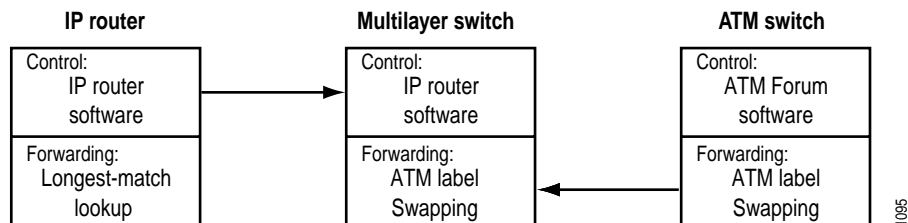
- ▶ IP Switching designed by Ipsilon/Nokia
- ▶ Tag Switching developed by Cisco Systems
- ▶ Aggregate Route-Based IP Switching (ARIS) designed by IBM Corporation
- ▶ IP Navigator delivered by Cascade/Ascend/Lucent Technologies
- ▶ Cell Switching Router (CSR) developed by Toshiba

Although these approaches had a number of characteristics in common, they were not interoperable because each relied on different technologies to combine IP routing and ATM switching into an integrated solution. However, by early 1997, many in the Internet community were impressed with the simplicity and elegance of these solutions that they began to view multilayer switching as the next logical evolutionary step for the design of large ISP backbone networks.

Similarities Among the Multilayer Switching Solutions

Each of the multilayer switching solutions sought to combine the best properties of IP routing and ATM switching, while still maintaining an IP focus. The fundamental approach adopted by these strategies was to take the control software from an IP router, integrate it with the forwarding performance of a label-swapping ATM switch, and create an extremely fast and cost efficient IP router (see Figure 4).

Figure 4: Multilayer Switch as a Fast IP Router



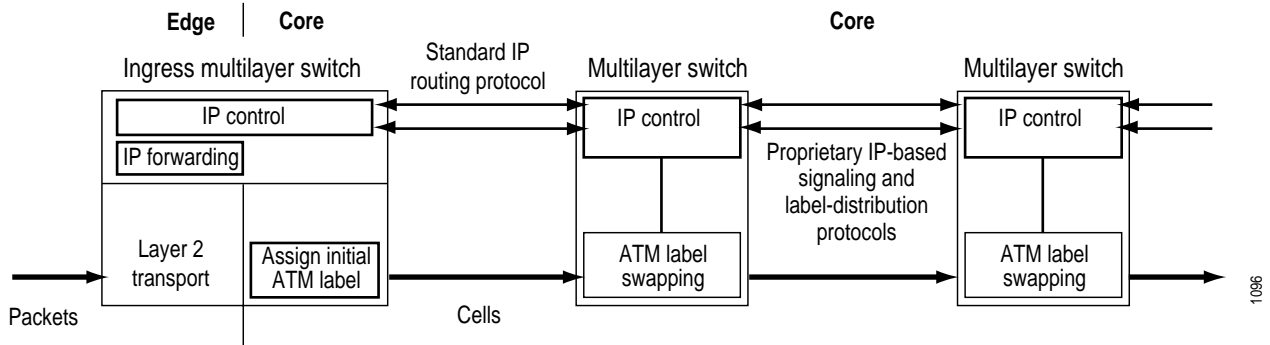
1095

For the control component, each multilayer switch ran standard IP routing software (OSPF, IS-IS, and BGP-4) and a proprietary label-binding mechanism. The routing software permitted multilayer switches to exchange Layer 3 network reachability information. The label-binding mechanism mapped Layer 3 routes to labels (that is, to ATM VPI/VCI) and distributed them to neighbors to establish LSPs across the core of the network. Running routing protocols on core systems rather than just edge systems provided a number of benefits that enhanced network operation:

- Eliminated the IP-over-ATM model's "n-squared" PVC scaling problem
- Reduced Interior Gateway Protocol stress by dramatically decreasing the number of peers that each router had to maintain
- Permitted information about the core's actual physical topology to be made available to Network layer routing procedures

For the forwarding component, multilayer switches used conventional ATM switching hardware and label swapping to forward cells across the core of the network (see Figure 5). However, the control procedures that assigned the labels to routes, distributed the labels among multilayer switches, and created the forwarding tables were managed by proprietary IP-based protocols, not ATM Forum protocols. ATM label swapping in the core of the network provided a number of benefits:

- Label swapping optimized network performance by leveraging the benefits of hardware-based forwarding. Vendors believed that this would promote the creation of a new generation of products that offered superior price-performance ratios while dramatically reducing the time to market. While this was a critical issue in the mid-1990s, it is not one of the lasting benefits of multilayer switching because technological advances have provided better approaches to building Internet backbone routers.
- Label swapping made explicit routing practical. An explicit route is a preconfigured sequence of hops that describes the path that traffic should take across a service provider's network, thus permitting the construction of a forwarding path that is different from the one typically created by destination-based routing. Explicit paths provide ISPs precise control over traffic flows, making it possible to support traffic engineering, QoS, and loop prevention.
- Label swapping provided an instrument to extend control beyond the limitations of conventional destination-based routing. Multilayer switching's ability to provide enhanced forwarding control beyond that supported by traditional routing mechanisms is its lasting contribution to network design. Later in this paper we examine how multilayer switching facilitates the deployment of new types of routing functionality.

Figure 5: Multilayer Switching Solutions—Integrating IP Routing and ATM Switching


By excluding the ATM Forum's routing and signaling protocols, multilayer switching reduced operational complexity by eliminating the need to coordinate and map between two different protocol architectures—IP and ATM. Although multilayer switching still used standard ATM VPI/VCI as labels, they were assigned and distributed using proprietary IP-based protocols rather than the standard ATM Forum protocols. This streamlined the integration of IP routing with ATM switching by eliminating the need to perform complex mapping between architectures. However, a critical limitation of the majority of multilayer switching solutions was that they were restricted to running over a cell-based ATM infrastructure, when the Internet was becoming increasingly packet-oriented.

Fundamental Differences Among the Multilayer Switching Solutions

While the various multilayer switching solutions had numerous features in common, they relied on two fundamentally different approaches to initiate the assignment and distribution of label bindings to establish LSPs:

- ▶ Data-driven model
- ▶ Control-driven model

Data-Driven Model

In the data-driven model, label bindings are created when user data packets arrive. A flow is a sequence of packets that have the same source and destination IP addresses and TCP or UDP port numbers. A multilayer switch can either create a label binding as soon as it sees the first packet in a traffic flow or wait until it has seen a number of packets in the flow. The benefit of waiting for a number of packets ensures that the flow is long enough to merit the overhead of assigning and distributing a label. Multilayer switching solutions that implemented the data-driven approach were IP Switching (Ipsilon) and the Cell Switching Router (Toshiba). Note that MPLS does not support the data-driven model.

The advantage of the data-driven model is that a label binding is created only when there is a traffic flow that uses the label binding. However, this model has a number of limitations for deployment in the core of a large ISP network, where there can be an enormous number of individual traffic flows:

- ▶ Each multilayer switch must provide sophisticated and high-performance packet classification capabilities to identify traffic flows.

- ▶ Typically, there is a latency between the recognition of a flow and the assignment of a label to the flow. This means that each multilayer switch must also support longest-match IP forwarding during the setup phase so packets that have not been assigned to a flow can be forwarded and not dropped.
- ▶ The amount of control traffic needed to distribute label bindings is directly proportional to the number of traffic flows.
- ▶ The presence of a significant number of relatively short-lived flows can impose a heavy burden on network operations.

Conventional wisdom dictates that the data-driven model does not have the scaling properties required for application in the core of the Internet.

Control-Driven Model

In the control-driven model, label bindings are created when control information arrives. Labels are assigned in response to the normal processing of routing protocol traffic, control traffic such as RSVP traffic, or in response to static configuration. Multilayer switching solutions that implemented the control-driven model were Tag Switching (Cisco Systems), IP Navigator (Ascend/Lucent), and ARIS (IBM). In addition, MPLS uses the control-driven model.

The control-driven model has a number of benefits for deployment in the core of a large ISP network:

- ▶ Labels are assigned and distributed before the arrival of user data traffic. This means that if a route exists in the IP forwarding table, a label has already been allocated for the route, so traffic arriving at a multilayer switch can be label swapped immediately.
- ▶ Scalability is significantly better than in the data-driven model, because the number of label switched paths is proportional to the number of entries in the IP forwarding table, not to the number of individual traffic flows. For traffic engineering in large ISP networks, scaling could be even better—proportional to the number of exit points in the network. Label assignment based on prefixes, rather than individual flows, permits a single label to represent a highly aggregated FEC.
- ▶ In a stable topology, the label assignment and distribution overhead is lower than in the data-driven model because label-switched paths are established only after a topology change or the arrival of control traffic, not with the arrival of each “new” traffic flow.
- ▶ Every packet in a flow is label switched, not just the tail-end of the flow as in the data-driven model.

Fundamental Problem with Multilayer Switching Solutions

Each multilayer switching solution maintained the IP control component and used ATM label swapping as the forwarding component. The challenge facing the ISP community was that each solution was proprietary and therefore not interoperable. Also, the majority of multilayer switching solutions required an ATM transport because they could not operate over mixed media infrastructures (Frame Relay, PPP, SONET, and LANs). If multilayer switching was to be widely deployed by ISPs, there had to be a multivendor standard that could run over any link layer technology. In early 1997, the IETF established the MPLS working group to produce a unified and interoperable multilayer switching standard.

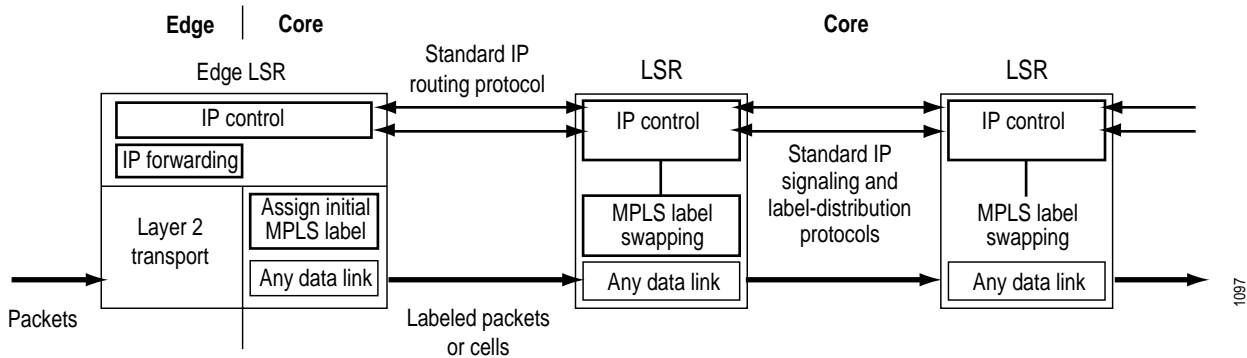
Multiprotocol Label Switching

MPLS is the latest step in the evolution of multilayer switching in the Internet. It is an IETF standards-based approach built on the efforts of the various proprietary multilayer switching solutions.

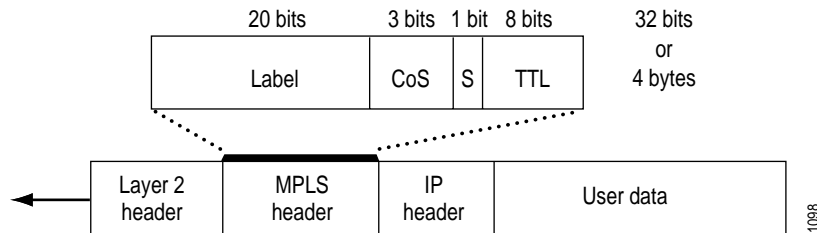
MPLS uses the control-driven model to initiate the assignment and distribution of label bindings for the establishment of label-switched paths (LSPs). LSPs are simplex in nature (traffic flows in one direction from the head-end toward the tail-end), duplex traffic requires two LSPs, one LSP to carry traffic in each direction. An LSP is created by concatenating one or more label switched hops, allowing a packet to be forwarded from one label-switching router (LSR) to another LSR across the MPLS domain. An LSR is a router that supports MPLS-based forwarding.

The MPLS control component centers around IP functionality, which is similar to proprietary multilayer switching solutions (see Figure 6). However, MPLS defines new standard-based IP signaling and label distribution protocols, as well as extensions to existing protocols, to support multivendor interoperability. MPLS does not implement any of the ATM Forum signaling or routing protocols so the complexity of coordinating two different protocol architectures is eliminated. In this way, MPLS brings significant benefits to a packet-oriented Internet.

Figure 6: Multiprotocol Label Switching



The MPLS forwarding component is based on the label-swapping algorithm. If the Layer 2 technology supports a label field (such as the ATM VPI/VCI or the Frame Relay DLCI fields), the native label field encapsulates the MPLS label. However, if the Layer 2 technology does not support a label field, the MPLS label is encapsulated in a standardized MPLS header that is inserted between the Layer 2 and IP headers (see Figure 7). The MPLS header permits any link layer technology to carry an MPLS label so it can benefit from label-swapping across an LSP.

Figure 7: MPLS Header

The 32-bit MPLS header contains the following fields:

- ▶ The label field (20-bits) carries the actual value of the MPLS label.
- ▶ The CoS field (3-bits) can affect the queuing and discard algorithms applied to the packet as it is transmitted through the network.
- ▶ The Stack (S) field (1-bit) supports a hierarchical label stack.
- ▶ The TTL (time-to-live) field (8-bits) provides conventional IP TTL functionality.

Requirements and Objectives

The charter of the MPLS working group is to standardize a base technology that combines the use of label swapping in the forwarding component with network layer routing in the control component. To achieve its objectives, the MPLS working group has to deliver a solution that satisfies a number of requirements, including:

- ▶ MPLS must run over any link layer technology, and just ATM.
- ▶ MPLS core technologies must support the forwarding of both unicast and multicast traffic flows.
- ▶ MPLS must be compatible with the IETF Integrated Services Model, including RSVP.
- ▶ MPLS must scale to support constant Internet growth.
- ▶ MPLS must support operations, administration, and maintenance facilities at least as extensive as those supported in current IP networks.

Common Misconceptions about MPLS

There are a number of misconceptions concerning the role of MPLS in the core of the Internet. Some in the Internet community believe that MPLS was developed to provide a standard that allowed vendors to transform ATM switches into high-performance Internet backbone routers. While this might have been one of the original goals of proprietary multilayer switching solutions in the mid-1990s, recent advances in silicon technology allow ASIC-based IP route lookup engines to run just as fast as MPLS or ATM VPI/VCI lookup engines. Although MPLS can enhance the forwarding performance of processor-based systems, accelerating packet forwarding performance was not the primary force behind the creation of the MPLS working group.

Others in the Internet community believe that MPLS was designed to completely eliminate the need for conventional, longest-match IP routing. This never was an objective of the MPLS working group because its members understood that traditional Layer 3 routing would always be required in the Internet.

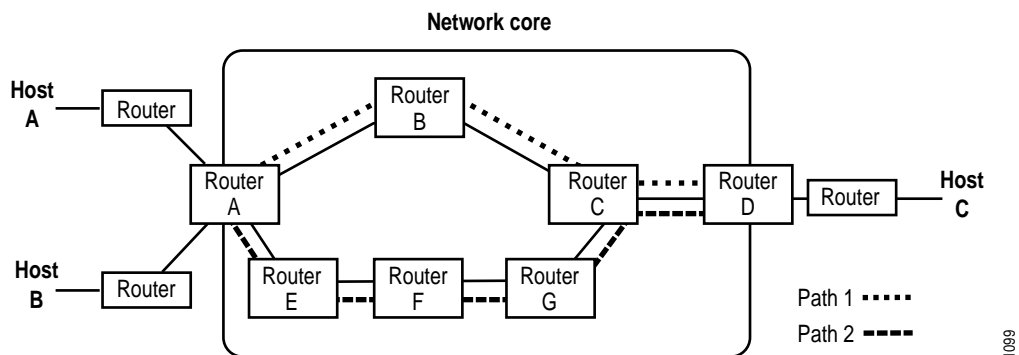
- ▶ Packet filtering at firewalls and ISP boundaries is a fundamental component of supporting security and enforcing administrative policy. Because packet filtering requires a detailed examination of packet headers, conventional Layer 3 forwarding is still required for these applications.
- ▶ It is unlikely that a large number of host systems will implement MPLS. This means that each packet transmitted by a host still needs to be forwarded to a first-hop Layer 3 device where the packet header can be examined prior to forwarding it towards its ultimate destination. The first-hop router can then either forward the packet using conventional longest-match routing or assign a label and forward the packet over an LSP.
- ▶ If a Layer 3 device along the path examines the IP header and assigns a label, the label represents an aggregate route because it is impossible to maintain label bindings for every host on the global Internet. This means that, at some point along the delivery path, the IP header must be examined by another Layer 3 device to determine a finer granularity to continue forwarding the packet. This router can elect to either forward the packet using conventional routing or assign a label and forward the packet over a new label switched path.
- ▶ At the last hop before the destination host, the packet must be forwarded using conventional Layer 3 routing because it is not practical to assign a separate label to every host on the destination subnetwork.

The Promise of MPLS

The question remains, “Why should an ISP consider deploying MPLS in the core of its network?” The most important benefit of MPLS is that it provides a foundation that permits ISPs to deliver new services that cannot be readily supported by conventional IP routing techniques. ISPs face the challenge of not only delivering superior baseline service, but also providing new services that distinguish them from their competition. MPLS allows service providers to control costs, provide better levels of base service, and offer new revenue-generating customer services.

Figure 8 illustrates how MPLS provides enhanced routing capabilities by supporting applications that require more than just destination-based forwarding. Assume that the routers in the core of the network perform conventional, longest-match IP forwarding. If either Host A or Host B transmits a packet to Host C, the packet follows Path 1 across the core of the network because this is the shortest path computed by the IGP.

Figure 8: MPLS Enhances Routing Functionality



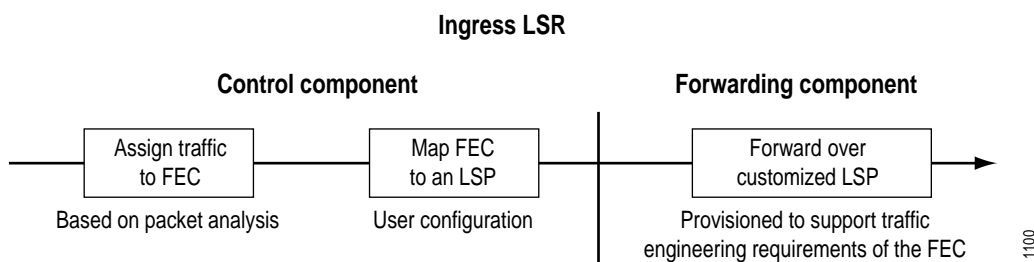
Suppose that the network administrator has been monitoring traffic statistics and needs to implement a policy to control congestion at Router B. The policy would reduce congestion at Router B by distributing the traffic load along different paths across the network. Traffic sourced by Host A and destined for Host C would follow the IGP shortest path, Path 1. Traffic sourced by Host B and destined for Host C would follow another path, Path 2. Using conventional IP routing, this policy cannot be implemented because all forwarding at Router A is based on the packet's destination address.

Now, if the routers in the core of the network function as LSRs, it is easy to implement a policy to reduce congestion at LSR B. The network administrator configures LSP 1 to follow Path 1. The network administrator configures LSP 2 to follow Path 2. Finally, the network administrator configures LSR A to put all traffic received from Host A and destined for Host C into LSP 1. Likewise, LSR A is configured to place all traffic received from Host B and destined for Host C into LSP 2. The ability to assign any FEC to a custom-tailored LSP gives the network administrator precise control of traffic as it flows through the provider's network.

With careful planning, MPLS provides ISPs an unprecedented level of control over traffic, resulting in a network that is more efficiently operated, supports more predictable service, and can offer the flexibility required to meet constantly changing customer expectations. You should note that the remainder of this section describes the potential for MPLS to assign traffic to FECs based on an extremely rich set of packet classification capabilities. Initial MPLS implementations will provide a more restricted set of packet classification capabilities which should be expected to evolve as the software implementing the control component matures.

As ISPs are required to roll out new customer services, the MPLS forwarding infrastructure can remain in place. New services can be deployed by simply modifying the control component that assigns packets to an FEC and then maps each FEC to a custom-built LSP (see Figure 9). For example, packets can be assigned to an FEC based on a combination of the destination subnetwork and application type, a combination of the source and destination subnetworks, a specific QoS requirement, an IP multicast group, or a Virtual Private Network (VPN) identifier. Similarly, network administrators can provision LSPs to satisfy specific FEC requirements—minimize the number of hops, meet specific bandwidth requirements, force traffic across certain links in the network, and so forth. The final step in evolving routing functionality is to configure the head-end LSR to place packets assigned to a particular FEC into an LSP that has been customized to support the FEC's requirements.

Figure 9: How MPLS Enhances Routing Functionality



MPLS Applications

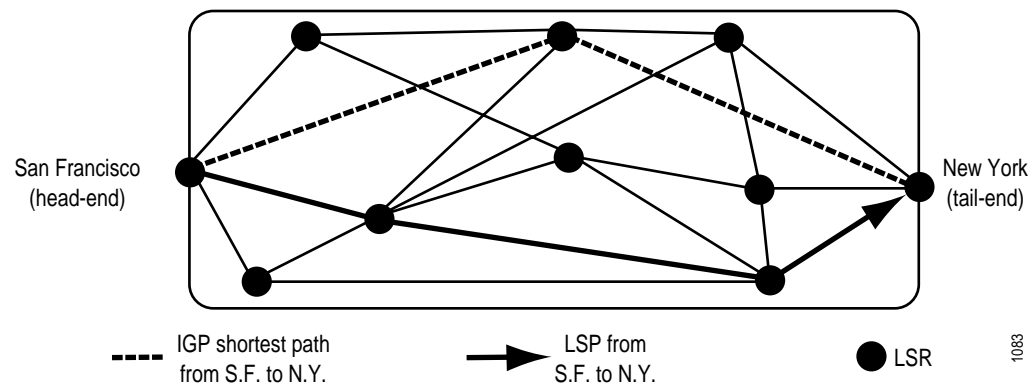
Currently there are three popular applications for MPLS in the core of large ISP networks:

- ▶ Traffic Engineering
- ▶ Class of Service (CoS)
- ▶ Virtual Private Networks (VPNs)

Traffic Engineering

Traffic engineering allows ISPs to move traffic flows away from the shortest path calculated by the IGP and onto potentially less congested physical paths across the network (see Figure 10). Traffic engineering is currently the primary application for MPLS because of the unprecedented growth in demand for network resources, the mission-critical nature of IP applications, and the increasingly competitive nature of the service provider marketplace. A successful traffic engineering solution can balance a network's aggregate traffic load on the various links, routers, and switches in the network so that none of its individual components is overutilized or underutilized. This results in a network that is more efficiently operated and provides more predictable service.

Figure 10: Traffic Engineering LSP vs. IGP Shortest Path across a Service Provider's Network



MPLS is well suited to provide the foundation to enable traffic engineering in large ISP networks for the following reasons:

- ▶ Support for explicit paths allows network administrators to specify the exact physical path that an LSP takes across the service provider's network.
- ▶ Per-LSP statistics can be used as input to network planning and analysis tools to identify bottlenecks and trunk utilization, and to plan for future expansion.
- ▶ Constraint-based routing provides enhanced capabilities that allow an LSP to meet specific performance requirements before it is established.
- ▶ An MPLS-based solution can run over packet-oriented networks and is not limited to ATM infrastructures.

Class of Service

MPLS can benefit service providers as they begin to roll out support for differentiated services (DiffServ). The differentiated services model defines a variety of mechanisms for classifying traffic into a small number of service classes. Subscribers are motivated to use the Internet as a public transport for a number of different applications ranging from traditional file transfer to delay sensitive services such as voice and video. To meet customer requirements, ISPs need to adopt not only traffic engineering techniques but also traffic classification technologies.

An ISP can take two approaches to support MPLS-based class of service forwarding:

- ▶ Traffic flowing through a particular LSP can be queued for transmission on each LSR's outbound interface based on the setting of the precedence bits carried in the MPLS header.
- ▶ An ISP can provision multiple LSPs between each pair of edge LSRs. Each LSP can be traffic engineered to provide different performance and bandwidth guarantees. The head end LSR could place high-priority traffic in one LSP, medium-priority traffic in another LSP, best-effort traffic in a third LSP, and less-than-best-effort traffic in a fourth LSP.

MPLS offers an ISP tremendous flexibility in the different types of services that it can provide its customers. The precedence bits are used only to classify packets into one of several classes of service. It is the ISP that determines the specific type of service that is supported by each service classification.

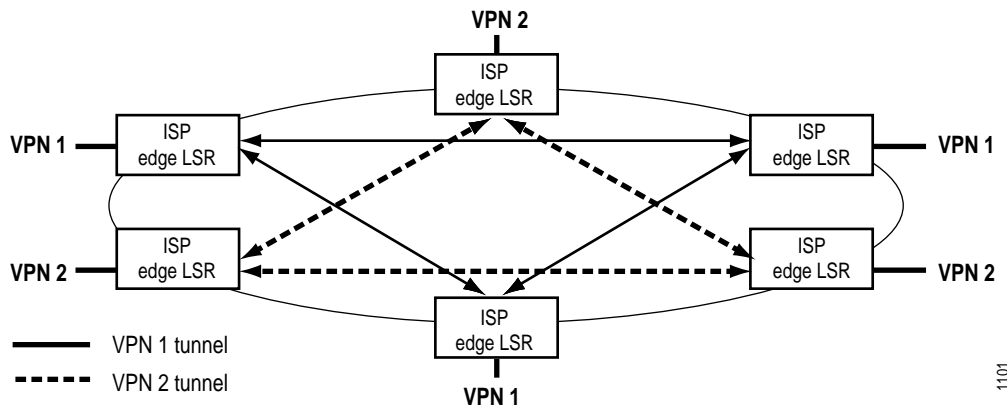
Virtual Private Networks

A virtual private network (VPN) simulates the operation of a private wide area network (WAN) over the public Internet. To offer a viable VPN service for its customers, an ISP must solve the problems of data privacy and support the use of non-unique, private IP addresses within a VPN. MPLS provides a simple and efficient solution to both of these challenges because it makes forwarding decisions based on the value of the label, not the destination address in the packet header.

VPNs are typically constructed using four fundamental building blocks:

- ▶ Firewalls to protect each customer site and provide a secure interface to the Internet
- ▶ Authentication to verify that each customer site exchanges data with only validated remote sites
- ▶ Encryption to protect data from examination or manipulation as it is transported across the Internet
- ▶ Tunneling encapsulation to provide a multiprotocol transport service and enable the use of the private IP address space within a VPN

MPLS allows ISPs to offer VPN services by providing a simple, flexible, and powerful tunneling mechanism (see Figure 11). An ISP can deploy a VPN by provisioning a set of LSPs to provide connectivity among the different sites in the VPN. Each VPN site then advertises to the ISP a set of prefixes that are reachable within the local site. The ISP's routing system distributes this information by piggybacking labels in routing protocol updates or by using a label distribution protocol. VPN Identifiers allow a single routing system to support multiple VPNs whose internal address spaces overlap with each other. Finally, each ingress LSR places traffic into LSPs based on a combination of a packet's destination address and VPN membership information.

Figure 11: MPLS Facilitates the Deployment of VPNs


Conclusion

All ISPs face the challenge of managing rapid growth. At the core of the network, WDM and greater fiber availability, as well as growth in subscriber demand, provide the incentive to carry Internet traffic over dedicated circuits or wavelengths. At the same time, the multiplexing, traffic engineering, and performance benefits gained from the IP-over-ATM model must be maintained in multiservice environments. MPLS provides the traffic engineering benefits of the IP-over-ATM model with some major advantages:

- Simpler network design and operation
- Improved scalability

MPLS is the latest step in the evolution of multilayer switching technology for the core of the Internet. It is an IETF standards-based approach built on the lessons learned from a number of proprietary multilayer switching solutions. MPLS combines the use of label swapping in the forwarding component with IP routing and standards-based IP signaling and label-distribution protocols in the control component. Furthermore, MPLS was specifically designed to run over any link-layer technology—not just an ATM infrastructure—facilitating the migration to the next-generation optical Internet based on SONET/WDM and IP/WDM infrastructures.

One of the most important benefits of MPLS is that it permits ISPs to deliver new services that cannot be readily supported by conventional IP routing techniques. MPLS offers enhanced routing capabilities by supporting more than just destination-based forwarding. Some of the new cost-reduction and revenue-generating services that can be deployed with MPLS include traffic engineering, CoS-based forwarding, and VPNs. By separating the control component from the forwarding component, MPLS provides the flexibility to evolve control functionality without changing the forwarding mechanism, thus uniquely positioning MPLS to support the deployment of enhanced forwarding capabilities that will be needed for the Internet to continue its explosive growth.

Related Reading

For related information, see the following white papers on <http://www.juniper.net>.

- ▶ *Traffic Engineering for the New Public Network*
- ▶ *RSVP Signaling Extensions for MPLS Traffic Engineering*

References

Textbooks

Davie, B., P. Doolan, and Y. Rekhter, *Switching in IP Networks: IP Switching, Tag Switching, and Related Technologies*, Morgan Kaufmann, 1998, ISBN 1-55860-505-3.

Metz, Christopher, *IP Switching: Protocols and Architectures*, McGraw-Hill, New York, 1999, ISBN 0-07-041953-1.

Requests for Comments

RFC 1953, *Ipsilon Flow Management Protocol Specification for IPv4 Version 1.0*, P. Newman, W.L. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, and G. Minshall, May 1996.

RFC 1954, *Transmission of Flow Labelled IPv4 on ATM Data Links Ipsilon Version 1.0*, P. Newman, W.L. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, and G. Minshall, May 1996.

RFC 2105, *Cisco Systems' Tag Switching Architecture Overview*, Y. Rekhter, B. Davie, D. Katz, E. Rosen, and G. Swallow, February 1997.

RFC 2297, *Ipsilon's General Switch Management Protocol Specification Version 2.0*, P. Newman, W. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, and G. Minshall, March 1998.

Internet Drafts

Baker, F., and Y. Rekhter, *Tag Switching with RSVP*, draft-baker-tags-rsvp-00.txt, December 1996.

Callon, R., G. Swallow, N. Feldman, A. Viswanathan, P. Doolan, and A. Fredette, *A Framework for Multiprotocol Label Switching*, draft-ietf-mpls-framework-02.txt, November 1997.

Callon, R., A. Viswanathan, and E. Rosen, *Multiprotocol Label Switching Architecture*, draft-ietf-mpls-arch-02.txt, July 1998.

Davie, B., Y. Rekhter, A. Viswanathan, S. Blake, V. Srinivasan, and E. Rosen, *Use of Label Switching With RSVP*, draft-ietf-mpls-rsvp-00.txt, March 1998.

Davie, B., T. Li, Y. Rekhter, and E. Rosen, *Explicit Route Support in MPLS*, draft-davie-mpls-explicit-routes-00.txt, November 1997.

Doolan, P., B. Davie, D. Katz, Y. Rekhter, and E. Rosen, *Tag Distribution Protocol*, draft-doolan-tdp-spec-01.txt, May 1997.

Feldman N., and A. Viswanathan, *ARIS Specification*, draft-feldman-aris-spec-00.txt, March 1997.

Viswanathan, A., N. Feldman, R. Boivie, and R. Woundy, *ARIS: Aggregate Route-Based IP Switching*, draft-viswanathan-aris-overview-00.txt, March 1997.

Next Generation Networks 1998: Conference Tutorials and Sessions

Awduche, Daniel, UUNET Technologies, Inc., "MPLS and Traffic Engineering in Service Provider Networks," Multilayer Switching Symposium, November 6, 1998.

Callon, Ross, Iron Bridge Networks, "Technologies for the Core of the Internet," Conference Tutorial, November 2, 1998.

Downey, Tom, Cisco Systems, Inc., "MPLS for Network Service Providers," Multilayer Switching Symposium, November 6, 1998.

Halpern, Joel, Newbridge Networks, "Technologies for Advanced Internet Services," Conference Tutorial, November 2, 1998.

Heinanen, Juha, Telia Finland, Inc., "MPLS—ATM Like Functionality for Router Backbones," Multilayer Switching Symposium, November 6, 1998.

McQuillan, John, McQuillan Consulting, "Major Trends in Broadband Networking," Conference Introduction, November 3, 1998.

Metz, Chris, Cisco Systems, Inc., "A Survey of Advanced Internet Protocols," Conference Tutorial, November 2, 1998.

Metz, Chris, Cisco Systems, Inc., "IP Switching in the Enterprise," Multilayer Switching Symposium, November 6, 1998.

Rybczynski Tony, Nortel Multimedia Networks, "Multilayer Switching in Enterprise WANs and VPNs," Multilayer Switching Symposium, November 6, 1998.

Sindhu, Pradeep, Juniper Networks, Inc., "Foundation for the Optical Internet," Next Generation Core Routers Session, November 3, 1998.

UUNET IW-MPLS '98 Conference

Barnes, Bill, UUNET Technologies, Inc., "Traffic Engineering with the Overlay Model," Traffic Engineering Session, November 13, 1998.

Huitema, Christian, Bellcore, "Scaling Issues in MPL Networks," Architecture and System Design Session, November 12, 1998.

St. Arnaud, Bill, Canarie, Inc., "MPLS and Architectural Issues for an Optical Internet," Architecture and System Design Session, November 12, 1998.

Swallow, George, Cisco Systems, Inc., "Traffic Engineering in MPLS Domains," Traffic Engineering Session, November 13, 1998.