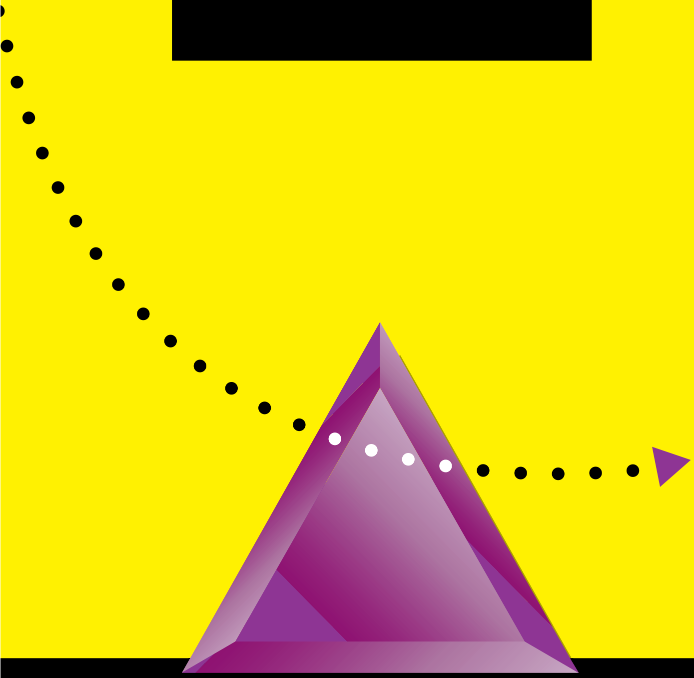
The Technology Guide Series

techguide.com

Multiprotocol Label Switching (MPLS)



This guide has been sponsored by

Ennovate Networks, Inc.

Visit ATG's Web Site

to read, download, and print

all the Technology Guides

in this series.

http://www.techguide.com

"The significant problems we face cannot be solved by the same level of thinking that created them."

Albert Einstein

Table of Contents

About the Editor…

Jerry Ryan is the vice president of Editorial Development for the Technology Guides on Communications and Networking. Mr. Ryan is also a principal at ATG. Mr. Ryan has developed and taught many courses in network analysis and design for carriers, government agencies and private industry. He has provided consulting support in the area of WAN and LAN network design, negotiation with carriers for contract pricing and services, technology acquisition, customized software development for network administration, billing and auditing of telecommunication expenses, project management, and RFP generation. He was the president and founder of Connections Telecommunications, Inc., a Massachusetts based company specializing in consulting, education, and software tools which address network design and billing issues. Mr. Ryan is a member of the Networld+Interop Program Committee. He holds a B.S. degree in electrical engineering.

Multiprotocol Label Switching (MPLS) was originally presented as a way of improving the forwarding speed of routers but is now emerging as a crucial standard technology that offers new capabilities for large scale IP networks. Traffic engineering, the ability of network operators to dictate the path that traffic takes through their network, and Virtual Private Network support are examples of two key applications where MPLS is superior to any currently available IP technology.

Although MPLS was conceived as being independent of Layer 2, much of the excitement generated by MPLS revolves around its promise to provide a more effective means of deploying IP networks across ATM-based WAN backbones. The Internet Engineering Task Force is developing MPLS with draft standards expected by the end of 1998. MPLS is viewed by some as one of the most important network developments of the 1990's. This Technology Guide will explain why MPLS is generating such interest.

The essence of MPLS is the generation of a short fixed-length 'label' that acts as a shorthand representation of an IP packet's header. This is much the same way as a ZIP code is shorthand for the house, street and city in a postal address, and the use of that label to make forwarding decisions about the packet. IP packets have a field in their 'header' that contains the address to which the packet is to be routed. Traditional routed networks process this information at every router in a packet's path through the network (hop by hop routing).

In MPLS, the IP packets are 'encapsulated' with these labels by the first MPLS device they encounter as they enter the network. The MPLS edge router analyses the contents of the IP header and selects an appropriate label with which to encapsulate the packet. Part of the great power of MPLS comes from the fact that, in contrast to conventional IP routing, this analysis can be based on more than just the destination address carried in the IP header. At all the subsequent nodes within the network the MPLS label, and not the IP header, is used to make the forwarding decision for the packet. Finally, as MPLS labeled packets leave the network, another edge router removes the labels.

In MPLS terminology, the packet handling nodes or routers are called Label Switched Routers (LSRs). The derivation of the term should be obvious; MPLS routers *forward packets by making* switching *decisions based on the MPLS* label*. This illustrates another of the key concepts in MPLS. Conventional IP routers contain 'routing tables' which are 'looked up' using the IP header from a packet to decide how to forward that packet. These tables are built by IP routing protocols (e.g., RIP or OSPF) which carry around IP reachability information in the form of IP addresses. In practice, we find that forwarding (IP header lookup) and control planes (generation of the routing tables) are tightly coupled. Since MPLS forwarding is based on labels it is possible to cleanly separate the (label-based) forwarding plane from the routing protocol control plane. By separating the two, each can be modified independently. With such a separation, we don't need to change the forwarding machinery, for example, to migrate a new routing strategy into the network.*

There are two broad categories of LSR. At the edge of the network, we require high performance packet classifiers that can apply (and remove) the requisite labels: we call these MPLS edge routers. Core LSRs need to be capable of processing the labeled packets at extremely high bandwidths.

This Technology Guide examines MPLS and the opportunities it offers to users and also to the service providers who are designing and engineering the next generation of IP networks. It also describes why new carrier-class edge devices will become a key component in the provisioning of future network services.

Introduction

Even though the standards are still in draft form, Multiprotocol Label Switching (or MPLS, as it is usually abbreviated) has become a technology that is key to the future of large-scale IP networks. MPLS has applications in the deployment of IP networks across ATM-based wide area networks, in providing traffic engineering capabilities to packet-based networks, in providing IP QoS capabilities, and in aiding the deployment of IP-based Virtual Private Networks (VPNs). These advances are critical to success for providers of the multiservice, multi-user, carrier-class internetworks that are now on the drawing boards.

MPLS is significantly different from the hop-by-hop processing methods of traditional networks. A short, fixed-length, easily-processable 'label' provides a shorthand representation of an IP packet's header in much the same way as a ZIP code is shorthand for the house, street and city in a postal address. Several manufacturers had developed proprietary solutions based on the label concept, which prompted the Internet Engineering Task Force (IETF) to begin the development of an interoperable standard to be called MPLS. In this Guide, MPLS refers to the IETF standards and label switching is used as a general reference to any label-based forwarding technique including MPLS.

This Technology Guide examines MPLS (at its current state of development) and describes why it was invented, what it does, what advantages it provides and where it appears to be headed. MPLS standards offer the promise of important new internetworking functionality; these are identified and discussed. The underlying protocols mechanisms are introduced and their relation to traditional routing explained. Finally, this Guide explains how new carrier-class edge switches will fit into MPLS-based IP network designs.

The Beginnings of MPLS

The TCP/IP protocol suite (and especially the IP protocol itself) is now the foundation for many public (the Internet) and private (the corporate Intranet) data networks. The forthcoming convergence of voice, data, and multimedia networks is also expected to be based largely on IP-based protocols, leading to the need for technical and operational improvements. Label switching is one of the industry's responses to this challenge.

Improving the original TCP/IP architecture, not only to differentiate among vendor products but also to create integrated public networks, has become a significant industry incentive. For example, IP networks need to evolve to support real-time packet delivery, integration of IP with ATM protocols, virtual public networks, and much larger size public networks. The number of hosts that can be attached, the number of routes that are possible and the bandwidth that is available all need to be highly scalable. Efficiency enhancements that improve switching price/performance and lower overall costs (which could stimulate the use of voice over IP, for example) are also eagerly anticipated. Using label switching for QoS support and providing features for explicit traffic engineering are viewed as part of the solution.

Label switching solutions can be characterized by their use of label swapping packet forwarding combined with IP control protocols and a label distribution mechanism. It is the differences in the details that distinguish among the techniques that have been proposed.

Although label switching tries to solve a wider range of problems than just the integration of IP and ATM, the difficulties associated with mapping between IP and ATM protocol models was a significant driver for the development of label switching technology. Over the last five years, a number of companies have attempted to blend the high-speed operation of ATM-based switching with the routing processes of the Internet's IP-based network layer. Four of these are noteworthy:

a) The *Cell Switching Router* (CSR) approach was developed by Toshiba and presented to the IETF in 1994. It was one of the earliest public proposals for using IP protocols to control an ATM switching fabric. CSR is designed to function as a router for connecting logical IP subnets in a classical 'IP over ATM' environment. Label switching devices communicate over standard ATM virtual circuits. CSR labeling is data-driven (i.e., labels are assigned on the basis of flows that are locally identified). The Flow Attribute Notification Protocol (FANP) is used to identify the dedicated VCs between CSR's and to establish the association between individual flows and individual dedicated VCs. The objective of the CSR is to allow 'cut through' forwarding of flows, i.e., to switch the ATM cell flow that constitutes the packet rather than reassembling it and making an IP level forwarding decision on it. CSRs have been deployed in commercial and academic networks in Japan.

b) *IP Switching*, developed by Ipsilon (who are now part of Nokia), was announced in early 1996 and has been delivered in commercial products. IP Switching enables a device with the performance of an ATM switch to act as a router, thereby overcoming the limited packet throughput of traditional routers. The basic goal of IP Switching is to integrate ATM switches and IP routing in a simple and efficient way (by eliminating the ATM control plane). IP Switching uses the presence of data traffic to drive the establishment of a label. A label binding protocol (called the Ipsilon Flow Management Protocol or IFMP) and a switch management protocol (called General Switch Management Protocol or GSMP) are defined. GSMP is used solely to control an ATM switch and the virtual circuits made across it.

c) *Tag Switching* is the label switching approach developed by Cisco Systems. In contrast to CSR and IP Switching, Tag Switching is a control-driven technique that does not depend on the flow of data to stimulate setting up of label forwarding tables in the router. A Tag Switching network consists of Tag Edge Routers and Tag Switching Routers, with packet tagging being the responsibility of the edge router. Standard IP routing protocols are used to determine the next hop for traffic. Tags are 'bound' to routes in a routing table and distributed to peers via a Tag Distribution Protocol. Tag switching is available on a number of products from Cisco.

d) *Aggregate Route-based IP Switching* (ARIS), IBM's label switching approach, is similar architecturally to Tag Switching. ARIS binds labels to aggregate routes (groups of address prefixes) rather than to flows (unlike CSR or IP Switching). Label bindings and label switched paths are set up in response to control traffic (such as routing updates) rather than data flows, with the egress router generally the initiator. Routers that are ARIS-capable are called Integrated Switch Routers. ARIS was designed with a focus on ATM as the Data Link Layer of choice (it provides loop prevention mechanisms that are not available in ATM). The ARIS Protocol is a peer-to-peer protocol that runs between ISRs directly over IP and provides a means to establish neighbors and to exchange label bindings. A key concept in ARIS is the "egress identifier". Label distribution begins at the egress router and propagates in an orderly fashion towards the ingress router.

Since multiple proprietary solutions for label-based switching is clearly not an acceptable direction, it was recognized that standards were needed and that an IETF Working Group had to be formed. A charter was agreed to in the IETF in early 1997 and the inaugural meeting of the working group was held in April 1997. After much deliberation, the term Multiprotocol Label Switching (MPLS) was selected as the 'vendor independent' name for the set of standards that will be produced.

The Internet Draft MPLS Framework states that the goal of standardization is to "integrate the label swapping forwarding paradigm with network layer routing" with an initial focus on IPv4 and IPv6. MPLS provides the mechanisms and these can be applied in various ways according to the network's needs.

Draft standards are not expected until the end of 1998, although vendors are already working on implementations. Those who build large MPLS-based IP networks and fully exploit the benefits of MPLS can be expected to become leaders in the next wave of internetwork expansion.

Challenges to Contemporary Networks

Enterprise network designers today face requirements that were just dreams when IP was first defined in the 1970's. Contemporary networks are being asked to support higher and higher volumes of best-effort data in the traditional Internet way (using file transfers, electronic mail, and WWW access); they are also being asked to differentiate among various classes of traffic that may include voice, music, and video. Quality of service has become a rallying cry for those who visualize a global convergence towards IP for all forms of communications.

The capabilities of the underlying network elements - the routers and switches that implement the protocols - have become critical to the ability to make progress towards this vision. However, many experts now believe that traditional hop-by-hop processing is beginning to reach its technological limit, and that a "paradigm shift" is needed in the forwarding process. The challenge is to evolve the IP network architecture in a way that simultaneously prepares for next generation networks, allows a smooth transition from the current environment, controls costs, and provides entrepreneurial opportunities for users and suppliers.

It has often been assumed that there was just one factor to consider - production of bigger, faster, cheaper routers. The explosive growth of the Internet and its projected expansion to many millions of IP addresses has put raw performance in the spotlight (and router manufacturers have responded with high capacity traditional routers). Label switching technology development, however, is being driven by much more than just the need for speed. Two of the most significant aspects are that:

- Different classes of traffic require specific service characteristics that must be guaranteed across the complete path through the network (and often across multiple autonomous systems). MPLS allows the creation of Label Switched Paths with different service characteristics.
- Carrier-class, multi-customer IP infrastructures require robust networks that can manage resources more effectively.

From the carriers' perspective, the efficient utilization of expensive network assets is the key to profitability. The traffic engineering capabilities of MPLS allow carriers a degree of control over the network's behavior that conventional IP technologies do not. From their customers perspective the bottom line is better service – the absence of congestion, for example.

Contemporary networks face major challenges in the following areas:

a) *Functionality*. Label switching provides new functions that were either unavailable or inefficient with conventional routing. Explicit routing to select a specific route that may not be the shortest route, is one example. Choosing a route on the basis of attributes other than the destination address, such as QoS, are also needed.

b) *Scalability*. Future networks need to be virtually unlimited in size. Routing information grows very quickly as the network grows, and can eventually overload a router by itself. Current techniques of overlaying IP routed networks on top of ATM or frame relay virtual circuits exacerbates this problem. MPLS requires the L2 devices (ATM switches for example) to be capable of running the IP control plane which ameliorates this problem. Traffic engineering, in the sense that it allows more efficient use of network resources also helps with 'scaling' the network.

c) *Evolvability*. One of the greatest challenges will be enabling change and growth without major network disruptions. Deterministic services need to be overlaid onto a non-deterministic IP network, multiple IP traffic types need to be accepted, and virtual private networks need to be established and removed. While the core of the network must increase in switching capacity, much of the evolution is driven by the edge device - the vendor/user demarcation point. A carrier-class device that incorporates new IP capabilities into an industry standard model is essential.

d) *Integration*. Application convergence for IP telephony is one example of systems integration and the overlay of the IP network on an ATM carrier infrastructure successfully is an example of network integration. Integration at all levels is a design requirement for an effective network.

MPLS Protocols and Functions

Routing and Switching Concepts

Several basic concepts that apply to any switching technology need to be reviewed prior to describing how MPLS works.

a) *Routing* is a term loosely used to describe the actions taken by the network to move packets through it. We speak of packets being 'routed' from 'a' to 'b', or of them being routed through a network or internetwork. There may be many routers in a network connected in some arbitrary fashion. Packets progress through the network by being sent from one machine to another toward their destination. Routing protocols (e.g. RIP, OSPF) enable each machine to understand which other machine is the 'next hop' that a packet should take toward its destination. Routers use the routing protocols to construct routing tables. When they receive a packet and have to make a forwarding decision, the routers 'look up' the routing table using the destination IP address in the packet as an index, thereby obtaining the identity of the 'next hop' machine. The construction of the tables and their use for look ups at forwarding time are essentially separate logical operations. Figure 1 illustrates these functions as they might occur in a router.

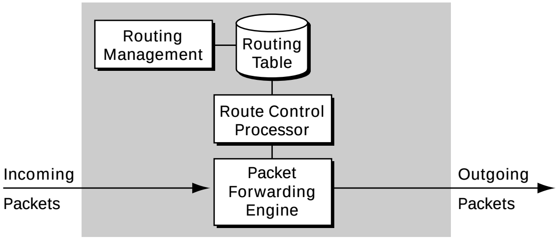


Figure 1

b) *Switching* is generally used to describe the transfer of data from an input to an output port of a machine where the selection of the output port is based on Layer 2 (e.g., ATM VPI/VCI) information.

c) The *control component* builds and maintains a forwarding table for the node to use. It works with the control components of other nodes to distribute routing information consistently and accurately, and also ensures that consistent local procedures are used to create the forwarding tables. Standard routing protocols (e.g., OSPF, BGP, and RIP) are used to exchange routing information among the control components. The control component must react when network changes occur (such as a link failure) but is not involved in the processing of individual packets.

d) The *forwarding component* performs the actual packet forwarding. It uses information from the forwarding table (as maintained by the router); information that is carried by the packet itself and a set of local procedures in order to make forwarding decisions. In a conventional router, a longest match algorithm compares the destination address in the packet with entries in the forwarding table until it obtains the 'best' available match. More importantly, the full decision-making process has to be repeated at each node along the path from source to destination. In an LSR, an (exact match) label swapping algorithm uses the label in the packet and a label-based forwarding table to obtain a 'new' label and output interface for the packet.

e) A *forwarding table* is the set of entries in a table that provides information to help the forwarding component perform its switching function. The forwarding table must associate each packet with an entry (traditionally the destination address) that provides instructions on where the packet is to go next.

f) A *Forwarding Equivalence Class (FEC)* is defined as any group of packets that can be treated in an equivalent manner for purposes of forwarding. An example of an FEC is the set of unicast packets whose destination addresses match a particular IP address prefix. Another FEC is the set of packets whose source and destination addresses are the same. FECs can be defined at different levels of granularity (for example, all packets matching a given address prefix is a coarser granularity than all packets from a given source going to a specific destination application port). Figure 2 illustrates the idea of FEC granularity.

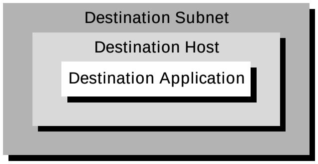


Figure 2

g) A *label* is a relatively short, fixed-length, unstructured identifier that can be used to assist in the forwarding process. Labels are associated with an FEC through a binding process. Labels are normally local to a single data link and have no global significance (as would an address). Labels are analogous to the DLCIs used in a Frame Relay network or the VPI/VCIs used in an ATM environment. Since ATM is a technology that already uses short fixed length fields to make switching decisions, label switching is believed to be an effective way of deploying IP over ATM. Labels are bound to an FEC (and therefore become meaningful) as a result of some event that indicates a need for the binding. These events can be divided into two categories:

- *Data-driven bindings* occur when traffic begins to flow, is submitted to the LSR and is recognized as a candidate for label switching. Label bindings are established only when needed, resulting in fewer entries in the forwarding table. Labels are assigned to individual IP traffic flows and not single packets. In an ATM network, this can result in the use of a substantial number of virtual circuits, which may limit network scalability.
- *Control-driven bindings* are established as a result of control plane activity and are independent of the data. Label bindings might be established in response to routing updates or receipt of RSVP messages. Control-driven label binding scales better than the data driven approach and for this reason is used in MPLS.

Label Switching

Label switching is an advanced form of packet forwarding that replaces conventional longest address match forwarding with a more efficient label swapping algorithm. There are three important distinctions between label switching and conventional routing:

	Conventional Routing	Label Switching
Full IP Header Analysis	Occurs at every node	Occurs only once at the network edge when label is assigned
Unicast & Multicast support	Requires multiple complex forwarding algorithms	One forwarding algorithm required
Routing decisions	Based on address only	Can be based on any number of parameters, such as QoS, VPN membership

A Label Switching Router is any device that supports both the standard IP control component (i.e., routing protocols, RSVP, etc.) and a label swapping forwarding component. Figure 3 shows a simple label switching network and illustrates the Edge LSRs (providing the ingress and egress functions) and Core LSRs (providing high speed switching). A label switching network serves the same purpose as any conventional routed network: it delivers traffic to one or more destinations. The addition of label-based forwarding complements conventional routing but does not replace it.

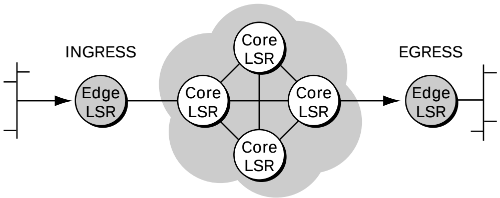


Figure 3

The Label Switching Forwarding Component

A label can be associated with a packet in several ways. Some networks can embed the label in the Data Link Layer header (the ATM VCI/VPI, and the Frame Relay DLCI specifically). The other option is to squeeze it into a small label header that sits between the Data Link header and the Data Link protocol-data-units (i.e., in between the Layer 2 header and the Layer 3 data being carried). These techniques allow label switching to be supported by virtually any Data Link including Ethernet, FDDI, and point-to-point links.

At the boundary of an MPLS network the edge LSRs make classification and forwarding decisions by examining the IP header in the unlabelled packets. The result is that appropriate labels are applied to the packets and they are then forwarded to an LSR that serves as the next hop toward the ultimate destination.

The LSR-generated, fixed-length "label" acts as a shorthand representation for the IP packet's header, thereby reducing the processing complexity at all subsequent nodes in the path. The label is generated during header processing at the LSR node. All subsequent nodes in the network use the label for their forwarding decisions. Of course the value of the label may, and usually does, change at each LSR in the path through the network. This is label switching after all!

As packets emerge from the core of an MPLS network, the edge LSRs that find they have to forward packets onto an unlabelled interface simply remove any label encapsulation before doing so.

When a core LSR receives a labeled packet, the label is first extracted and it is used as an index into the forwarding table that resides in the LSR. When the entry indexed by the incoming label is found, the outgoing label is extracted and added to the packet and the packet is then sent out the outgoing interface(s) to the next hop(s) that are specified in the entry (multicast involves multiple outgoing packets). Label switching forwarding tables may be implemented at the node level (a single table per node) or at the interface level (one table per interface).

What is most important about label-based forwarding is that only a single forwarding algorithm is needed for all types of switching and this can be implemented in hardware for extra speed.

The Label Switching Control Component

Labels are attached to the packets by an 'upstream' LSR. The 'downstream' LSR that receives these labeled packets must know (or find out) what to do with them. It is the responsibility of the label switching control component to handle this task. It uses the contents of an entry in the label switching forwarding table as its guide.

Needless to say, establishment and maintenance of table entries are essential functions and must be performed by each LSR. The label switching control component is responsible for distributing routing information among the LSRs in a consistent fashion and for executing the procedures that are used by the LSRs to convert this information into a forwarding table.

The label switching control component includes all the conventional routing protocols (e.g., OSPF, BGP, PIM, and so on). These routing protocols provide the LSRs with the mapping between the FEC and the next hop addresses. In addition, the LSR must:

- Create the bindings between the labels and the FECs
- Distribute those bindings to other LSRs
- Construct its own label forwarding table

The binding between a label and an FEC can be data-driven (i.e., be the result of the presence of specific types of traffic flow) or can be control-driven (i.e., be directed by the topology as represented in routing updates or other control messages).

Each of these binding techniques have numerous options. The decision to establish labeled flow can be based on multiple criteria (i.e., the source of the data may indicate a lot of data is to be expected). Data-driven label binding establishes active label bindings only when there is an immediate need (i.e., traffic has been presented for forwarding). Both topology changes and traffic changes must be distributed. Control-driven binding is based on management knowledge resulting from route processing and resource reservations. Although both techniques have been used, the emerging MPLS standards will be based on the control-driven model.

Distribution of Label Information

A label switching forwarding table entry provides, at a minimum, information about the outgoing interface and a new label, but may also contain other information. It might, for example, indicate the output queuing discipline to be applied to the packet. The incoming label uniquely identifies a single entry in this table.

Every label that is distributed must be bound to an entry in the forwarding table. This binding may be performed in the local LSR or be supplied by a remote LSR. The current version of MPLS uses downstream binding in which locally bound labels are used as incoming labels, and remotely bound labels are used as outgoing labels. It should be noted that the opposite of this, called upstream binding, is also feasible. For MPLS, the entries in the forwarding table are established as follows:

The ***Next Hop*** is provided by the routing protocols (the FEC to next hop mapping),
The ***Incoming Label*** is provided by creating a local binding between an FEC and the label, and
The ***Outgoing Label*** is provided by a remote binding between an FEC and the label.

The MPLS architecture uses both local control (the LSR can decide to create and advertise a binding without waiting to receive a binding from a neighbor for the same FEC) and egress control (the LSR waits for a binding from its downstream neighbor before allocating a label and advertising it upstream).

Knowledge of the bindings between locally chosen labels and the FECs they are associated with must be disseminated to adjacent LSRs for use in creating their own forwarding tables. The information in the forwarding table must also track changes in the network in a consistent fashion. Afterall, it is the label on the incoming packet that is used to discover the rules for forwarding the packet.

Label information can be distributed in two ways:

a) Piggybacking on a Routing Protocol

MPLS label binding information may be added to conventional routing protocols for distribution although only control-driven schemes can support this method. Piggybacking on the normal operation of routing protocols ensures consistency of the forwarding information and avoids the need for yet another protocol. Unfortunately, not all subnets use routing and not all routing protocols are easily able to handle labels so this is not a complete answer for label distribution.

b) Use of a Label Distribution Protocol

Following the Cisco TDP model from Tag Switching, the MPLS working group has embarked on the definition of a new protocol specifically for the distribution of label binding information called the Label Distribution Protocol (LDP). The LDP can be used for both control- and data-driven schemes. The disadvantage of an explicit LDP is that it adds complexity (yet another new protocol has to be supported) and its use needs to be coordinated with the operation of its associated routing protocols.

The definition of the LDP for use with MPLS is an ongoing effort and a number of the details have not yet been completed. It is anticipated that the working group will be able to converge on a stable definition of a Version 1.0 LDP by the end of 1998.

The Role of the Edge LSR

It is the responsibility of the edge LSRs to classify traffic and apply and remove labels to and from packets. As has been noted previously, labels can be assigned on the basis of factors other than destination address. The edge LSR determines whether the traffic is a long-lasting flow, implements management policies and access controls, and performs aggregation of traffic into larger flows when possible. These are all functions that need to be performed at the boundary between the IP and MPLS worlds. Thus, the capabilities of edge LSRs will be key to the success of an overall label switching environment. It is also a point of control and management for the service provider. We expect to see a new generation of products specifically designed as MPLS edge routers.

This new generation of edge LSRs will have the following capabilities:

- *Wirespeed IP flow classification capabilities*: This will allow these products to assign QoS values and apply labels to IP flows without any degradation in forwarding performance; and
- *Extensive VPN capabilities*: To take advantage of MPLS when provisioning VPNs, these products must be able to run multiple forwarding tables so that VPN customers can be separated within the LSR.

Benefits and Advantages of MPLS

One of the major advantages of MPLS is the fact that it will be a standards-based implementation of label switching technology. The development of standards results in an open environment with multiple manufacturers' products all being interoperable. Competition also results in lower prices, leads to more innovative features and stimulates early availability. MPLS is expected to have broad industry support and will eventually supplant the current proprietary solutions.

The real questions to be asked are: *What are the benefits and advantages of using label switching? Is label switching a necessary step in the evolution of the TCP/IP architecture? Would improvements to conventional routing meet the perceived application requirements?*

a) *Explicit Routes*

A key feature of MPLS is its support for explicit routes. Explicitly routed Label Switched Paths are far more efficient than the source route option in IP. They also provide some of the functionality needed for traffic engineering. Explicitly routed paths also have attractions as 'opaque tunnels' where they can carry any type of traffic (e.g. SNA, IPX) that the two cooperating tunnel end points agree on. Because the intermediate LSRs that 'carry' the tunnel see only the MPLS labels arbitrary traffic can be carried in packets sent on the tunnel.

b) *Virtual Private Networks (VPNs)*

Many organizations use private networks built using leased lines to connect multiple sites. A carrier offering that emulates the secure, reliable, and predictable behavior of these networks over shared carrier facilities holds the promise of providing extra service revenues to the carrier, while also lowering the cost of ownership borne by the customer. VPNs are an emulation of these Private Networks across carrier facilities in such a manner that each customer perceives himself to be running on a Private Network. The carrier's infrastructure has been 'Virtualized' to support many independent mutually invisible networks. MPLS is a key ingredient in building such networks; the MPLS labels can be used to isolate traffic between (and even within) VPNs.

c) *Multiprotocol and Multilink Support*

The label switching forwarding component is not specific to a particular Network Layer. For example, the same forwarding component could be used when doing label switching with IP as well as with IPX. Label switching is also able to operate over virtually any Data Link Layer protocols, although the initial emphasis is on ATM. The 'Multi' in MPLS applies above and below the label switching layer!

d) *Evolvability*

Label switching also has the advantage of a clean separation between its control and forwarding functions. Each part can evolve without impacting the other part, which makes the evolution of networks easier, less costly, and less prone to errors.

e) *Inter-domain Routing*

Label switching provides a more complete separation between inter- and intra-domain routing. This improves the scalability of routing processes and, in fact, reduces the route knowledge required within a domain. This is a benefit to ISPs and carriers who may have a large amount of transit traffic (i.e., traffic whose source and destination is not on the network).

f) *Support for All Traffic Types*

One other advantage of label switching which is not generally visible to the user is that it supports all types of forwarding: unicast, unicast with type of service, and multicast packets.

Label switching also improves upon the various methods that have been tried for integrating IP with ATM-based subnetworks. This may remove the need for complex procedures and protocols that deal with issues such as address resolution and the different models for multicast and resource reservation.

Label switching can be used with QoS attributes that, in turn, allow different classes of ISP access service to be defined ("first-class" vs. "coach-class" for example).

Label switching can permit the actual IP header in a packet to be encrypted since all that must be available to the LSR is the label itself. In this way the sources and destinations of the data are no longer observable while in transit.

Summary

Multiprotocol Label Switching is destined to provide a new technical foundation for the next generation of multi-user, multiservice internetworks. The promise is for higher performance, another order of magnitude increase in scalability, improved and expanded functionality, and the flexibility to match the user's quality of service requirements more closely. While the expansion of the Internet has been a major driver for development of label switching, it is not the only, or even the most important, factor.

Label switching provides significant improvements in the packet forwarding process by simplifying the processing, avoiding the need to duplicate header processing at every step in the path, and creating an environment that can support controlled QoS. Several vendor-specific solutions exist today and IETF MPLS standards are expected within a year. Deployment of MPLS allows a closer integration of IP and ATM, supports service convergence, and offers new opportunities for traffic engineering and VPN support..

By adding fixed size labels to packet flows, the way we add ZIP codes to mail to help with sorting, packet processing performance can be improved, QoS controls can be more easily applied and very large global public networks can be built. All of this results in better networks with more functions at lower cost. MPLS is a new technology that is just beginning to be recognized as beneficial. The basic standards will soon be completed and products will be delivered quickly afterward. It is fully expected that MPLS will see widespread deployment in both public and private IP networks, paving the way for true convergence of telephony, video, and computing services.

CASE STUDY: Application 1 – Enabling IP over ATM

Transporting IP over an ATM network creates scalability, network performance, and network administration concerns. The topology as shown if Figure 1 creates a large number of router adjacencies that result in a less than optimal performance of routing protocols. This also requires the set-up and administration of a large number of control ATM VCs which become cumbersome to support and maintain. To create a fully meshed network, each router has to be joined to each other router via an ATM VC. This creates the need for N(N-1)/2 virtual circuits (where N equals the number of nodes). This network topology does not scale well, creating explosive growth in the number of virtual circuits as the network gets larger.

Problem of IP over ATM

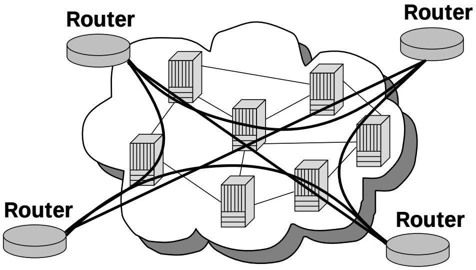


Figure 1

IP over ATM in a MPLS Network

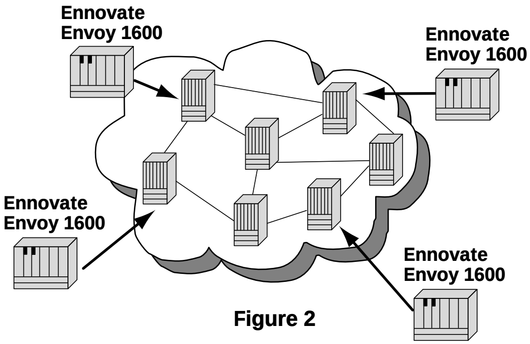


Figure 2

In Figure 2, the MPLS network has two new classes of devices. A label edge router (LER) such as the Ennovate Envoy 1600 and Label Switching Router (LSRs). The LER as the name implies, is positioned at the edge of the service providers' networks. The LER devices are responsible for IP flow classification and label imposition. The LSR devices located in the core are responsible for forwarding at Layer 2 while participating in the exchange of Layer 3 routing information. An LSR device could be an upgraded ATM switch.

This new network topology significantly reduces the number of routing adjacencies (the example in Figure 2 requires only one adjacency between the Ennovate Envoy 1600 and the closest LSR) and the need for establishing a mesh network of control VCs is eliminated. This result is less complexity and a lower cost of ownership.

The Ennovate Envoy 1600 allows for allocating of QoS to individual IP flows and then mapping these flows to the appropriate ATM class of service as shown if Figure 3. This is accomplished by the classification and marking of IP flows through Ennovate's advanced custom designed 'High Touch Routing' ASIC. The ASIC examines all elements of the IP and TCP/UDP header and makes critical classification and marking decisions based on one or several parameters at wire-line speeds.

Network Service Providers can now define innovative new IP services over *existing* ATM networks . As these services become successfully deployed, MPLS will allow these networks to scale and accommodate the increased network traffic.

Enables use of ATM QoS

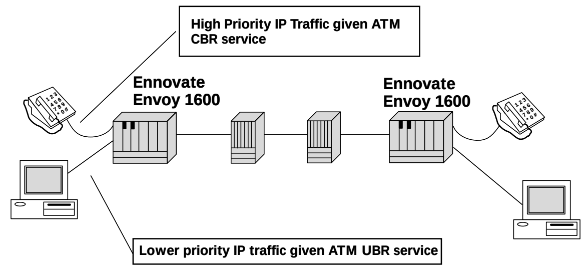


Figure 3

Application 2 – Traffic Engineering

Conventional dynamic routing was designed to be very resilient and self-healing in the advent of a network failure. Parameters such as hop count have been used to ensure the best path through the network. This was sufficient in a best effort delivery IP environment.

Dynamic Routing

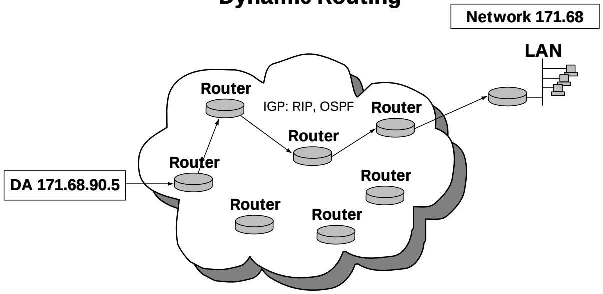


Figure 4

However, as Network Service Providers provision new services, best effort delivery is not sufficient. The need to engineer and control traffic patterns through the network is crucial to the network operator. MPLS provides for explicit routing. Explicit routing is the capability to direct traffic along a route other than the one that IP routing would choose. This is accomplished by establishing an explicitly routed Label Switched Path (LSP) through the network. This LSP can be thought of as an opaque tunnel which network traffic can be sent through. This traffic flows from the beginning to the end of the tunnel without the need for any direction from the devices along the LSP. This provides Network Service Providers an important tool allowing them to fully utilize important network assets (bandwidth/switches) and support new services. The Ennovate Envoy 1600 has the capabilities to iniate the setting up of these LSPs across an MPLS network and then to classify traffic so that it enters the appropriate LSP. These abilities will be key to the deployment of new premium IP services.

Traffic Engineering

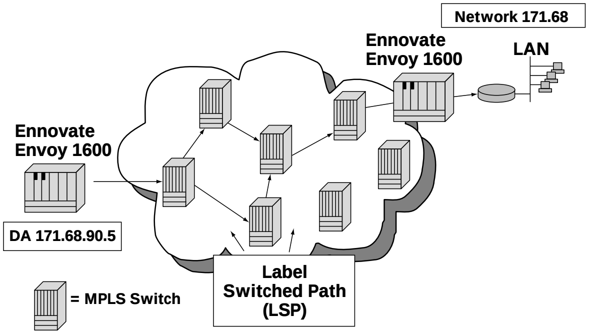


Figure 5

Application 3 – Virtual Private Networks (VPNs)

Analysts estimate the market for provisioned VPN services to be a multi-billion dollar service opportunity for the Network Service Providers. These services are anticipated to progressively replace existing private line and frame relay networks. Major corporations are expected to build mission-critical intranets and extranets using these new services. However, there are a number of issues addressed by MPLS that will help ensure the successful deployment of VPNs.

- Quality of service. As noted in the previous application, the capabilities to do explicit routing within MPLS helps the Network Service Provider engineer networks capable of sustaining quality of service.
- Many corporations do not have the globally unique IP addresses required for routing in a public network. The IPv4 architecture requires that the network part of the IP address must be unique (and routing in the network is based on this fact). MPLS helps by encapsulating a non-unique address in a unique (within the MPLS domain) label.

Using MPLS to Provision VPN's

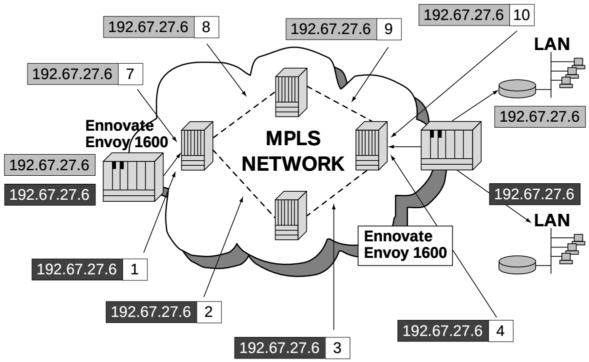


Figure 6

The Ennovate Envoy 1600 incorporates important features that enable the deployment of VPN services.

- Virtual Routers. A unique ‘Virtual Router’ technology solves the VPN private address problem by supporting multiple forwarding tables. These forwarding tables are used to keep each enterprise address space separate. Within the core of the network, these addresses can be kept distinct using ATM or frame VCs, IP tunnels or MPLS labels. A Network Service Provider can now utilize one router infrastructure to economically provision new services. VPN users can easily connect to the Internet, with integral Network Address Translation (NAT) via a separate forwarding table.

- Quality of Service. The “High Touch Routing” capabilities of the Ennovate Envoy 1600 allows Network Service Providers to support different qualities of services for VPNs. This is critical in the move to integrated multi-service VPNs where voice, video, and data require different levels of service.
- Secure VPN Membership Protocol. The Secure VPN Membership protocol provides the following capabilities:
  — Dynamically determines the set of nodes that are connected to various VPNs
  — Authentication to ensure VPN security
  — Dynamic creation of IP tunnels or other paths to create virtual links to interconnect VPNs
- Partitionable Network Management. Ennovate’s Network Management System allows for service provisioning management by Network Service Providers and for virtual network management by the corporate VPN manager.

The table below summarizes and contrasts an MPLS-based solution to a conventional router-based solution in each of the application areas described above.

	Conventional IP Network	**Ennovate Envoy 1600 in a MPLS Network**
Quality of Service	No differential IP QoS support	Maps specific IP flows to ATM Classes of Service
Traffic Engineering	Best Effort Delivery only	Label Switched Paths (LSPs) can be manually created through the network to ensure QoS guarantees and provision new services
VPN Support	One Router Network per Customer VPN Best Effort Routing for VPNs Static VPN creation	Virtual Routers provide separate routing tables per customer VPN Provides different QoS parameters for VPNs Secure VPN Membership protocol for authentication, dynamic path creation and dynamic node determination
Scalability	Creates large number of Router adjacencies which adversely effects routing protocol performance	Creates small number of adjacencies for optimal protocol routing performance
Voice and Data Integration	Voice over IP treated as best effort delivery	Standard voice quality achievable with Traffic Engineering and QoS support Built-in T1/E1 cross connect for smooth service migration of voice traffic
Administration	Cumbersome to set-up and support large number of VCs	Eliminates needs to create mesh of VCs

GLOSSARY

AAL—ATM Adaptation Layer. A protocol layer that allows higher layer protocols to run over ATM virtual circuits. AAL5, for example, enables segmentation and re-assembly of variable-length packets into cells on an ATM Virtual Circuit.

ARIS—Aggregate Route-based IP Switching. ARIS is IBM's label switching proposal and is similar architecturally to Tag Switching.

ATM—Asynchronous Transfer Mode. A high speed, switching transfer mode in which the information is organized into fixed cells to transmit data, voice, and video. It is asynchronous in the sense that the recurrence of cells containing information from an individual user is not necessarily periodic.

BGP—Border Gateway Protocol. An IP protocol used to exchange routing information between network domains.

CLEC—Competitive Local Exchange Carrier. A competitor to the local telephone companies that has been granted permission by the State Regulatory Commission to offer local telephone services. CLECs are sometimes called alternative local exchange carriers.

Control Component—A function performed by a router that builds and maintains a forwarding table and works with other control components of other nodes to distribute routing information.

CPE—Customer Premises (or Provided) Equipment. This is equipment such as telephone systems, modems, and terminals installed at the customer's site.

DLCI—Data Link Control Identifier. A label used in Frame Relay networks to identify specific frame relay virtual circuits.

CSR—Cell Switch Router. Toshiba's label switching technology.

Edge LSR—A carrier-class Label Switching Router located at the edge of the carrier network which first classifies IP flows and applies a label.

Egress Identifier—A concept used in ARIS, referring to the identifier of the last LSR in a label switched path.

Explicit Routing—The ability to select a specific route not based on the shortest path and destination address, but based on a specific policy, quality of service, or virtual private network membership.

FANP—Flow Attribute Notification Protocol. The protocol used by CSRs to notify neighbors that a flow has been selected for switching.

FEC—Forwarding Equivalence Class. A group of packets treated identically when transported through a network.

Flow—A set of packets being transmitted between a set of hosts or a pair of transport protocol ports on a pair of hosts.

Flow Identifier—An object used by CSR, IP Switching, and other data-driven approaches to label a flow to be switched.

Forwarding—The process of transmitting a packet from a source to a destination on either a switch or router.

Forwarding Component—The forwarding process performed by a router to do the actual packet transport based on information contained in the routing table.

GSMP—General Switch Management Protocol. The protocol defined by Ipsilon to allow communication between an IP switch controller and an ATM switch.

IETF—Internet Engineering Task Force. The organization that provides the coordination of standards and specification development for TCP/IP networking.

IFMP—Ipsilon Flow Management Protocol. The label binding protocol which an IP Switch uses to notify its neighbors that a flow has been selected for label switching.

IP—Internet Protocol. A Layer 3 (network layer) protocol that contains addressing information and some control information that allows packets to be routed.

IP Flow Classification—A function performed by an edge LSR that categorizes IP traffic flows, assigns QoS values and associates labels with identified FECs.

IP Switching—First generation label switching technology developed by Ipsilon (now Nokia).

IPv6—Internet Protocol Version 6

ISP—Internet Service Provider. A company that provides Internet access services to individual users and businesses.

ISR—The ARIS term for a Label Switching Router.

IXC—Inter-Exchange Carrier. A public switching network carrier that provides (in conjunction with the local exchange carriers—LECs) interLATA access services.

Label—A short, fixed-length identifier that is used to determine the forwarding of a packet using the exact match algorithm and which is usually rewritten during forwarding.

Label Binding—An association between a label and a FEC which may be advertised to neighbors to establish a label switched path.

Label Switching—The generic term used here to describe all approaches to forwarding IP packets using a label swapping forwarding algorithm under the control of network layer routing algorithms.

LDP—Label Distribution Protocol. A new protocol being defined by the IETF designed to disseminate and track changes to locally assigned labels and the FECs they are associated with between adjacent LSRs.

LEC—Local Exchange Carrier. Any company authorized by the state public utility commission to sell local service.

Longest Match—The forwarding algorithm most often used for IP forwarding, in which a fixed-length IP address is compared against the variable-length entries in a routing table, looking for the entry that matches the most leading bits in the address.

LSR—Label Switching Router. A LSR is a device that supports both the standard IP control component (i.e. routing protocols, RSVP, etc) and a label swapping forwarding component.

MPLS—Multi-Protocol Label Switching. The name of the IETF working group that is standardizing label switching.

Multicast—Single packets copied to a specific subset of network addresses. These addresses are specified in the destination-address field. In contrast, in a broadcast, packets are sent to all devices in a network.

NSP—Network Service Provider

OC-n—Optical Carrier-n. An ITU-T-specified physical interface for transmission over optical fiber at n times the basic rate of 51.84 Mbps (e.g., OC-3 is at 155.52 Mbps).

OSPF—Open Shortest Path First. A standard link-state Internet Protocol (IP) routing protocol

QoS—Quality of Service. The capability to differentiate between traffic and service types so that one or more classes of traffic can be treated differently than other types.

PIM—Protocol Independent Multicast. A multicast routing protocol being standardized in the IETF.

Port—(1) A physical interface to a switch or router. (2) An identifier used by transport protocols to distinguish application flows between a pair of hosts.

RIP—Routing Information Protocol. A popular standard IP routing protocol.

Router—A layer 3 (Network Layer) device that maintains a forwarding table and forwards packets through a network.

Routing Domain—That part of a network that is controlled by a specific routing protocol.

RSVP—Resource Reservation Protocol. A protocol for reserving network resources to provide quality of service guarantees to application flows.

SVC—Switched Virtual Circuit. A connection between two end points used by a connection-oriented Layer 2 technology such as ATM or Frame Relay that can be dynamically switched through the network.

Switching—A general term given to the processing of a message, packet, cell, or frame. Most often is applied to Layer 2 – Data Link Control services.

Tag—Another name for a label, used in Cisco's Tag Switching.

Tag Edge Routers—Devices at the edge of the network that perform packet tagging in a Tag Switching Network.

Tag Switching Routers—Devices in the core of a Tag Switching network that switches tags assigned by Tag Edge Routers.

Tag Switching—Tag Switching is the label switching approached developed by Cisco Systems that has been submitted to the IETF for publication.

TCP—Transmission Control Protocol. The widely used reliable byte stream delivery protocol.

TFIB—Tag Forwarding Information Base. The data structure used in Tag Switching to hold information about incoming and outgoing tags and the associated FECs.

TOS—Type of Service.

UNI—User Network Interface. The interface, defined as a set of protocols and traffic characteristics, between the CPE (user) and the ATM network (ATM switch).

Unicast—Equivalent to point-to-point transmission.

VPI / VCI—Virtual Path Identifier / Virtual Channel Identifier. A field in the ATM header used to identify the virtual circuit to which a cell belongs.

VPN—Virtual Private Network. In a VPN, resources (such as bandwidth and buffer space) are provided, on-demand, to the users (usually by the public carriers) in such a way that the users view a certain partition of that network as a private network. The advantage of the VPNs, over the dedicated private networks, is lower cost and dynamic use of network resources.

WAN—Wide Area Network. This is a network that spans a large geographic area.

For further information contact:
Ennovate Networks, Inc.
330 Codman Hill Rd.
Boxborough, MA 01719
Phone: 978 263-2002
Fax: 978 263-1099
www.ennovatenetworks.com

NOTES

NOTES

NOTES

NOTES

NOTES

Visit ATG's Web Site
to read, download, and print
all the Technology Guides
in this series.

http://www.techguide.com

"The significant problems we face cannot be solved by the same level of thinking that created them."

Albert Einstein

This Technology Guide is one in a comprehensive series of Guides that provide objective information and practical guidance on technologies related to Communications & Networking, the Internet, Document Management, Data Warehousing, and Enterprise Solutions. Our team of technical editors writes each Technology Guide to assist IT and business professionals in making informed decisions about all aspects of technology application development and strategic deployment.

techguide.com is supported by a consortium of leading technology providers. Ennovate Networks, Inc. has lent its support to produce this Guide.

Visit our Web Site at www.techguide.com to view and print this Guide, as well as all of our other Technology Guides. This is available as a free service.

techguide.com