Multiprotocol Label Switching Overview

This chapter describes the Multiprotocol Label Switching (MPLS) distribution protocol. MPLS is a high-performance packet forwarding technology that integrates the performance and traffic management capabilities of data link layer (Layer 2) switching with the scalability, flexibility, and performance of network-layer (Layer 3) routing. It enables service providers to meet challenges brought about by explosive growth and provides the opportunity for differentiated services without necessitating the sacrifice of existing infrastructure.

The MPLS architecture is remarkable for its flexibility:

- Data can be transferred over any combination of Layer 2 technologies
- Support is offered for all Layer 3 protocols
- Scaling is possible well beyond anything offered in today's networks.

Specifically, MPLS can efficiently enable the delivery of IP services over an ATM switched network. It supports the creation of different routes between a source and a destination on a purely router-based Internet backbone. Service providers who use MPLS can save money and increase revenue and productivity.

Note Label switching on a router requires that Cisco Express Forwarding be enabled on that router. Refer to the Cisco Express Forwarding feature documentation for configuration information.

- MPLS Virtual Private Networks, page 9
- MPLS Quality of Service, page 9

Finding Feature Information

Your software release may not support all the features documented in this module. For the latest caveats and feature information, see Bug Search Tool and the release notes for your platform and software release. To find information about the features documented in this module, and to see a list of the releases in which each feature is supported, see the feature information table at the end of this module.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

MPLS Tag Switching Terminology

Beginning with Cisco IOS Release 12.1, the Tag Switching distribution protocol has been replaced with the MPLS distribution protocol. The Tag Switching command-line interface (CLI) commands are supported but will be discontinued in a future release.

The table below lists tag switching terms (found in earlier releases of this document) and the equivalent MPLS terms used in this document.

Table 1: Equivalency Table for Tag Switching and MPLS Terms

Old Tag Switching Terminology	New MPLS Terminology
Tag Switching	Multiprotocol Label Switching (MPLS)
Tag (short for Tag Switching)	MPLS
Tag (item or packet)	Label
TDP (Tag Distribution Protocol)	LDP (Label Distribution Protocol) Cisco TDP and LDP (MPLS Label Distribution Protocol) are nearly identical in function, but use incompatible message formats and some different procedures. Cisco is changing from TDP to a fully compliant LDP.
Tag Switched	Label Switched
TFIB (Tag Forwarding Information Base)	LFIB (Label Forwarding Information Base)
TSR (Tag Switching Router)	LSR (Label Switching Router)
TSC (Tag Switch Controller)	LSC (Label Switch Controller)

Old Tag Switching Terminology	New MPLS Terminology
ATM-TSR (ATM Tag Switch Router)	ATM-LSR (ATM Label Switch Router, such as the Cisco BPX 8650 switch)
TVC (Tag VC, Tag Virtual Circuit)	LVC (Label VC, Label Virtual Circuit)
TSP (Tag Switch Path)	LSP (Label Switch Path)
XTag ATM (extended Tag ATM port)	XmplsATM (extended MPLS ATM port)

MPLS Commands and Saved Configurations

During the transition period from tag switching to MPLS, if a configuration command has both MPLS and tag switching forms, the tag switching version is written to saved configurations. For example, you can configure MPLS hop-by-hop forwarding for a router POS interface by issuing the following commands:

```
Router# configure terminal
Router(config)# interface POS3/0
Router(config-if)# mpls ip
```

In this example, the **mplsip** command has a tag switching form (**tag-switchingip**). After you enter these commands and save this configuration or display the running configuration by means of the **showrunningconfiguration**command, the configuration commands appear as follows:

```
interface POS3/0
tag-switching ip
```

Saving the tag switching form of commands (that have both tag switching and MPLS forms) allows for backward compatibility. You can use a new router software image to modify and write configurations, and then later use configurations created by the new image with earlier software versions that do not support the MPLS forms of commands

Using the tag switching forms of the commands allows older software that supports tag switching commands, but not new MPLS commands, to successfully interpret interface configurations.

MPLS Tag Switching CLI Command Summary

The table below summarizes general-purpose MPLS commands. Except where otherwise noted, these MPLS commands have been derived from existing tag-switching commands to preserve the familiar syntax of existing commands that formed the basis for implementing new MPLS functionality. The tag-switching versions of the command will be discontinued in a future release.

Table 2: Summary of MPLS Commands Described in this Document

Command	Corresponding Tag Switching Command	Description
debug mpls adjacency	**debug tag-switching adjacency**	Displays changes to label switching entries in the adjacency database.

Command	Corresponding Tag Switching Command	Description
debug mpls events	**debug tag-switching events**	Displays information about significant MPLS events.
debug mpls lfib cef	**debug tag-switching tfib cef**	Prints detailed information about label rewrites being created, resolved, and deactivated as CEF routes are added, changed, or removed.
debug mpls lfib enc	**debug tag-switching tfib enc**	Prints detailed information about label encapsulations while label rewrites are created or updated and placed into the label forwarding information base (LFIB).
debug mpls lfib lsp	**debug tag-switching tfib tsp**	Prints detailed information about label rewrites being created and deleted as TSP tunnels are added or removed.
debug mpls lfib state	**debug tag-switching tfib state**	Traces what happens when label switching is enabled or disabled.
debug mpls lfib struct	**debug tag-switching tfib struct**	Traces the allocation and freeing of LFIB-related data structures, such as the LFIB itself, label-rewrites, and label-info data.
debug mpls packets	**debug tag-switching packets**	Displays labeled packets switched by the host router.
interface atm	**interface atm**	Enters interface configuration mode, specifies ATM as the interface type, and enables the creation of a subinterface on the ATM interface.
mpls atm control-vc	**tag-switching atm control-vc**	Configures the VPI and VCI to be used for the initial link to the label switching peer device.
mpls atm vpi	**tag-switching atm vpi**	Configures the range of values to be used in the VPI field for label VCs.
mpls ip (global configuration)	**tag-switching ip (global configuration)**	Enables MPLS forwarding of IPv4 packets along normally routed paths for the platform.
mpls ip (interface configuration)	**tag-switching ip (interface configuration)**	Enables MPLS forwarding of IPv4 packets along normally routed paths for a particular interface.

Command	Corresponding Tag Switching Command	Description
mpls ip default-route	**tag-switching ip default-route**	Enables the distribution of labels associated with the IP default route.
mpls ip propagate-ttl	**tag-switching ip propagate-ttl**	Sets the time-to-live (TTL) value when an IP packet is encapsulated in MPLS.
mpls ip ttl-expiration pop	N/A	Forwards packets using the global IP routing table or the original label stack, depending on the number of labels in the packet.
mpls label range	**tag-switching tag-range downstream**	Configures the range of local labels available for use on packet interfaces. **Note** The syntax of this command differs slightly from its tag-switching counterpart.
mpls mtu	**tag-switching mtu**	Sets the per-interface maximum transmission unit (MTU) for labeled packets.
show mpls forwarding-table	**show tag-switching forwarding-table**	Displays the contents of the label forwarding information base (LFIB).
show mpls interfaces	**show tag-switching interfaces**	Displays information about one or more interfaces that have been configured for label switching.
show mpls label range	N/A	Displays the range of local labels available for use on packet interfaces.

Benefits

MPLS provides the following major benefits to service provider networks:

- Scalable support for Virtual Private Networks (VPNs)--MPLS enables VPN services to be supported in service provider networks, thereby greatly accelerating Internet growth.

The use of MPLS for VPNs provides an attractive alternative to the building of VPNs by means of either ATM or Frame Relay permanent virtual circuits (PVCs) or various forms of tunneling to interconnect routers at customer sites.

Unlike the PVC VPN model, the MPLS VPN model is highly scalable and can accommodate increasing numbers of sites and customers. The MPLS VPN model also supports “any-to-any” communication among VPN sites without requiring a full mesh of PVCs or the backhauling (suboptimal routing) of traffic across the service provider network. For each MPLS VPN user, the network of the service provider appears to function

as a private IP backbone over which the user can reach other sites within the VPN organization, but not the sites of any other VPN organization.

From a user perspective, the MPLS VPN model enables network routing to be dramatically simplified. For example, rather than needing to manage routing over a topologically complex virtual backbone composed of many PVCs, an MPLS VPN user can generally employ the backbone of the service provider as the default route in communicating with all of the other VPN sites.

- Explicit routing capabilities (also called constraint-based routing or traffic engineering)--Explicit routing employs "constraint-based routing," in which the path for a traffic flow is the shortest path that meets the resource requirements (constraints) of the traffic flow.

In MPLS traffic engineering, factors such as bandwidth requirements, media requirements, and the priority of one traffic flow versus another can be taken into account. These traffic engineering capabilities enable the administrator of a service provider network to perform the following tasks:

- • Control traffic flow in the network
  - Reduce congestion in the network
  - Make best use of network resources

Thus, the network administrator can specify the amount of traffic expected to flow between various points in the network (thereby establishing a traffic matrix), while relying on the routing system to perform the following tasks:

- • Calculate the best paths for network traffic
  - Set up the explicit paths to carry the traffic

Label Switching Functions

In conventional Layer 3 forwarding mechanisms, as a packet traverses the network, each router extracts all the information relevant to forwarding the packet from the Layer 3 header. This information is then used as an index for a routing table lookup to determine the next hop for the packet.

In the most common case, the only relevant field in the header is the destination address field, but in some cases, other header fields might also be relevant. As a result, the header analysis must be done independently at each router through which the packet passes. In addition, a complicated table lookup must also be done at each router.

In label switching, the analysis of the Layer 3 header is done only once. The Layer 3 header is then mapped into a fixed length, unstructured value called a *label* .

Many different headers can map to the same label, as long as those headers always result in the same choice of next hop. In effect, a label represents a *forwarding equivalence class* --that is, a set of packets which, however different they may be, are indistinguishable by the forwarding function.

The initial choice of a label need not be based exclusively on the contents of the Layer 3 packet header; for example, forwarding decisions at subsequent hops can also be based on routing policy.

Once a label is assigned, a short label header is added at the front of the Layer 3 packet. This header is carried across the network as part of the packet. At subsequent hops through each MPLS router in the network, labels are swapped and forwarding decisions are made by means of MPLS forwarding table lookup for the label

carried in the packet header. Hence, the packet header does not need to be reevaluated during packet transit through the network. Because the label is of fixed length and unstructured, the MPLS forwarding table lookup process is both straightforward and fast.

Distribution of Label Bindings

Each> label switching router (LSR) in the network makes an independent, local decision as to which label value to use to represent a forwarding equivalence class. This association is known as a label binding. Each LSR informs its neighbors of the label bindings it has made. This awareness of label bindings by neighboring routers is facilitated by the following protocols:

- Label Distribution Protocol (LDP)--enables peer LSRs in an MPLS network to exchange label binding information for supporting hop-by-hop forwarding in an MPLS network
- Tag Distribution Protocol (TDP)--Used to support MPLS forwarding along normally routed paths
- Resource Reservation Protocol (RSVP)--Used to support MPLS traffic engineering
- Border Gateway Protocol (BGP)--Used to support MPLS virtual private networks (VPNs)

When a labeled packet is being sent from LSR A to the neighboring LSR B, the label value carried by the IP packet is the label value that LSR B assigned to represent the forwarding equivalence class of the packet. Thus, the label value changes as the IP packet traverses the network.

MPLS and Routing

A label represents a forwarding equivalence class, but it does not represent a particular path through the network. In general, the path through the network continues to be chosen by the existing Layer 3 routing algorithms such as OSPF, Enhanced IGRP, and BGP. That is, at each hop when a label is looked up, the next hop chosen is determined by the dynamic routing algorithm.

MPLS Traffic Engineering

MPLS traffic engineering software enables an MPLS backbone to replicate and expand upon the traffic engineering capabilities of Layer 2 ATM and Frame Relay networks. MPLS is an integration of Layer 2 and Layer 3 technologies. By making traditional Layer 2 features available to Layer 3, MPLS enables traffic engineering. Thus, you can offer in a one-tier network what now can be achieved only by overlaying a Layer 3 network on a Layer 2 network.

Traffic engineering is essential for service provider and Internet service provider (ISP) backbones. Such backbones must support a high use of transmission capacity, and the networks must be very resilient so that they can withstand link or node failures.

MPLS traffic engineering provides an integrated approach to traffic engineering. With MPLS, traffic engineering capabilities are integrated into Layer 3, which optimizes the routing of IP traffic, given the constraints imposed by backbone capacity and topology.

Why Use MPLS Traffic Engineering

WAN connections are an expensive item in an ISP budget. Traffic engineering enables ISPs to route network traffic to offer the best service to their users in terms of throughput and delay. By making the service provider more efficient, traffic engineering reduces the cost of the network.

Currently, some ISPs base their services on an overlay model. In the overlay model, transmission facilities are managed by Layer 2 switching. The routers see only a fully meshed virtual topology, making most destinations appear one hop away. If you use the explicit Layer 2 transit layer, you can precisely control how traffic uses available bandwidth. However, the overlay model has numerous disadvantages. MPLS traffic engineering achieves the traffic engineering benefits of the overlay model without running a separate network, and without needing a nonscalable, full mesh of router interconnects.

How MPLS Traffic Engineering Works

MPLS traffic engineering automatically establishes and maintains LSPs across the backbone by using RSVP. The path that an LSP uses is determined by the LSP resource requirements and network resources, such as bandwidth.

Available resources are flooded by means of extensions to a link-state-based Interior Gateway Protocol (IGP).

Traffic engineering tunnels are calculated at the LSP head based on a fit between required and available resources (constraint-based routing). The IGP automatically routes the traffic onto these LSPs. Typically, a packet crossing the MPLS traffic engineering backbone travels on a single LSP that connects the ingress point to the egress point.

MPLS traffic engineering is built on the following Cisco IOS mechanisms:

- IP tunnel interfaces--From a Layer 2 standpoint, an MPLS tunnel interface represents the head of an LSP. It is configured with a set of resource requirements, such as bandwidth and media requirements, and priority.

From a Layer 3 standpoint, an LSP tunnel interface is the head-end of a unidirectional virtual link to the tunnel destination.

- MPLS traffic engineering path calculation module--This calculation module operates at the LSP head. The module determines a path to use for an LSP. The path calculation uses a link-state database containing flooded topology and resource information.
- RSVP with traffic engineering extensions--RSVP operates at each LSP hop and is used to signal and maintain LSPs based on the calculated path.
- MPLS traffic engineering link management module--This module operates at each LSP hop, does link call admission on the RSVP signalling messages, and does bookkeeping of topology and resource information to be flooded.
- Link-state IGP (Intermediate System-to-Intermediate System (IS-IS) or OSPF--each with traffic engineering extensions)--These IGPs are used to globally flood topology and resource information from the link management module.
- Enhancements to the SPF calculation used by the link-state IGP (IS-IS or OSPF)--The IGP automatically routes traffic onto the appropriate LSP tunnel based on tunnel destination. Static routes can also be used to direct traffic onto LSP tunnels.

- Label switching forwarding--This forwarding mechanism provides routers with a Layer 2-like ability to direct traffic across multiple hops of the LSP established by RSVP signalling.

One approach to engineering a backbone is to define a mesh of tunnels from every ingress device to every egress device. The MPLS traffic engineering path calculation and signalling modules determine the path taken by the LSPs for these tunnels, subject to resource availability and the dynamic state of the network. The IGP, operating at an ingress device, determines which traffic should go to which egress device, and steers that traffic into the tunnel from ingress to egress.

A flow from an ingress device to an egress device might be so large that it cannot fit over a single link, so it cannot be carried by a single tunnel. In this case, multiple tunnels between a given ingress and egress can be configured, and the flow is load-shared among them.

MPLS Virtual Private Networks

Using MPLS VPNs in a Cisco IOS network provide the capability to deploy and administer scalable Layer 3 VPN backbone services including applications, data hosting network commerce, and telephony services to business customers. A VPN is a secure IP-based network that shares resources on one or more physical networks. A VPN contains geographically dispersed sites that can communicate securely over a shared backbone.

A one-to-one relationship does not necessarily exist between customer sites and VPNs; a given site can be a member of multiple VPNs. However, a site can associate with only one VPN routing and forwarding instance (VRF). Each VPN is associated with one or more VPN VRFs. A VRF includes routing and forwarding tables and rules that define the VPN membership of customer devices attached to CE routers. A VRF consists of the following:

- IP routing table
- CEF table
- Set of interfaces that use the CEF forwarding table
- Set of rules and routing protocol parameters to control the information in the routing tables

VPN routing information is stored in the IP routing table and the CEF table for each VRF. A separate set of routing and CEF tables is maintained for each VRF. These tables prevent information from being forwarded outside a VPN and also prevent packets that are outside a VPN from being forwarded to a router within the VPN.

MPLS Quality of Service

The quality of service (QoS) feature for MPLS enables network administrators to provide differentiated types of service across an MPLS network. Differentiated service satisfies a range of requirements by supplying for each packet transmitted the particular kind of service specified for that packet by its QoS. Service can be specified in different ways, for example, using the IP precedence bit settings in IP packets.

In supplying differentiated service, MPLS QoS offers packet classification, congestion avoidance, and congestion management. The table below lists these functions and their descriptions.

Table 3: QoS Services and Features

Service	QoS Function	Description
Packet classification	Committed access rate (CAR). Packets are classified at the edge of the network before labels are assigned.	Classifies packets according to input or output transmission rates. Allows you to set the MPLS experimental bits or the IP Precedence or DSCP bits (whichever is appropriate).
Congestion avoidance	Weighted Random Early Detection (WRED). Packet classes are differentiated based on drop probability.	Monitors network traffic to prevent congestion by dropping packets based on the IP Precedence or DSCP bits or the MPLS experimental field.
Congestion management	Class-based weighted fair queueing (CBWFQ). Packet classes are differentiated based on bandwidth and bounded delay.	An automated scheduling system that uses a queueing algorithm to ensure bandwidth allocation to different classes of network traffic.

Note MPLS QoS lets you duplicate Cisco IOS IP QoS (Layer 3) features as closely as possible in MPLS devices, including label edge routers (LERs), LSRs, and ATM-LSRs. MPLS QoS functions map nearly one-for-one to IP QoS functions on all interface types.

For more information on configuration of the QoS functions (CAR, WRED, and CBWFQ), refer to the *Cisco IOS Quality of Service Solutions Configuration Guide* .

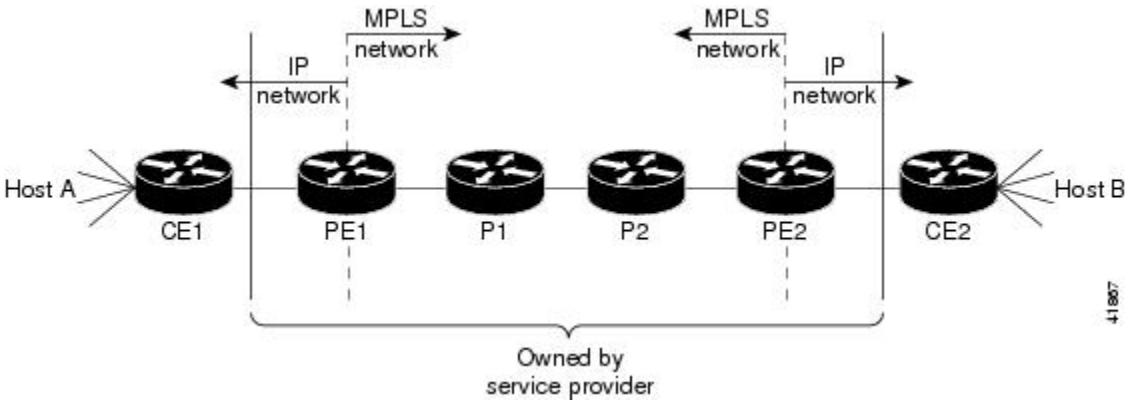
For complete command syntax information for CAR, WRED, and WFQ, refer to the *Cisco IOS Quality of Service Solutions Command Reference.*

Specifying the QoS in the IP Precedence Field

When you send IP packets from one site to another, the IP Precedence field (the first three bits of the DSCP field in the header of an IP packet) specifies the QoS. Based on the IP precedence marking, the packet is given the desired treatment such as the latency or the percent of bandwidth allowed for that quality of service. If the service provider network is an MPLS network, then the IP precedence bits are copied into the MPLS EXP field at the edge of the network. However, the service provider might want to set a QoS for a MPLS packet to a different value determined by the service offering.

This feature allows the service provider to set the MPLS experimental field instead of overwriting the value in the IP precedence field belonging to a customer. The IP header remains available for the customer's use; the QoS of an IP packet is not changed as the packet travels through the MPLS network.

The figure below shows an MPLS network that connects two sites of a IP network belonging to a customer.



Note The network is bidirectional, but for the purpose of this document the packets move left to right.

In the figure above, the symbols have the following meanings displayed in the table below:

Table 4: Device Symbols

Symbol	Meaning
CE1	Customer equipment 1
PE1	Service provider edge router (ingress LSR)
P1	Service provider router within the core of the network of the service provider
P2	Service provider router within the core of the network of the service provider
PE2	Service provider edge router (egress LSR)
CE2	Customer equipment 2

Note Notice that PE1 and PE2 are at the boundaries between the MPLS network and the IP network.

In the figure above, the following behavior occurs:

- Packets arrive as IP packets at PE1, the provider edge router (also known as the ingress label switching router).
- PE1 sends the packets as MPLS packets.
- Within the service provider network, there is *no IP Precedence field* for the queueing mechanism to look at because the packets are MPLS packets. The packets remain MPLS packets until they arrive at PE2, the provider edge router.

- PE2 removes the label from each packet and forwards the packets as IP packets.

This MPLS QoS enhancement allows service providers to classify packets according to their type, input interface, and other factors by setting (marking) each packet within the MPLS experimental field without changing the IP Precedence or DSCP field. For example, service providers can classify packets with or without considering the rate of the packets that PE1 receives. If the rate is a consideration, the service provider marks in-rate packets differently from out-of-rate packets.

Note The MPLS experimental bits allow you to specify the QoS for an MPLS packet. The IP Precedence/DSCP bits allow you to specify the QoS for an IP packet.