


# Quality of Service for Multi-Protocol Label Switching Networks



Is there a quality-of-service (QoS) architecture for MPLS? How does it differ from the IP QoS architectures?

A. MPLS doesn't define a new QoS architecture. Most of the work on MPLS QoS has focused on supporting current IP QoS architectures. There are two QoS architectures defined for IP: Integrated Services (IntServ) and Differentiated Services (DiffServ). IntServ defines per-flow QoS and uses RSVP as the signaling mechanism used by applications to request QoS from the network. MPLS can support per-flow QoS with the extensions made to RSVP to propagate bindings between flows and labels. The LABEL\_REQUEST and LABEL objects added to RSVP enable downstream label allocation using the PATH and RESV messages. These extensions are commonly used, however, for implementing resource reservation for flow aggregates in MPLS Traffic Engineering. They are not used for per-flow QoS due to the limited scalability that such approach would have in a service provider backbone. On the other hand, DiffServ defines a QoS architecture based on flow aggregates that requires traffic to be conditioned and marked at the network edges and internal nodes to give different QoS treatment to packets based on their markings. MPLS packets need to carry the packet marking in their headers because LSRs do not examine the IP header during forwarding. A 3-bit field in the MPLS shim header is used for this purpose. The DiffServ functionality of an LSR is almost identical to that provided by an IP router with respect to the QoS treatment given to packets (per-hop behavior in DiffServ terms).

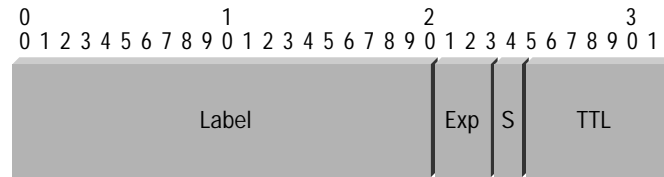
Have any QoS architectures been standardized for MPLS?

A. Cisco has proposed the specifications for MPLS support of the DiffServ architecture. Cisco has also been involved in MPLS enhancements for RSVP. No standardization has been proposed to date for MPLS support for IntServ. The latest version of the technical specifications can be found in the IETF MPLS work group home page at <http://www.ietf.org/html.charters/mpls-charter.html>

What are the experimental bits? Are they the same as the EXP bits or EXP field? How is this field used for implementing QoS?

A. The format of the MPLS shim header is defined in RFC 3032. That definition reserves 3 bits for experimental use. This field is commonly referred to as the EXP field (or simply as the EXP bits) and is used to support DiffServ on an MPLS network. These bits help define the QoS treatment (per-hop behavior) that

a node should give to a packet. It is the equivalent of the DiffServ Code Point (DSCP) in the IP network. A DSCP generally defines a class and in some cases a drop precedence within a class. The EXP bits are generally used to carry all the information encoded in the IP DSCP. In some cases, however, the EXP bits are used exclusively to encode the dropping precedence.



Label	Label Value, 20 Bits
Exp	Experimental Use, 3 Bits
S	Bottom of Stack, 1 Bit
TTL	Time to Live, 8 Bits

Why is this field called experimental? Is it appropriate to deploy a technology that uses a field that has been named experimental? Why are the EXP bits used for QoS?

A. The original Cisco proposal for Tag Switching (which later became MPLS) called that field the class of service (CoS) field. However, there was no consensus on the IETF for defining that 3-bit field as the CoS field. According to RFC 3032, that field is reserved for experimental use and as a result, it's now widely known as the EXP field or EXP bits. However, the most common use of those bits is for QoS purposes and the name of the field is not expected to change (they are used differently by GMPLS). The fact that this 3-bit field is called experimental does not mean that its use should be discouraged or considered arbitrary. Other MPLS vendors are also implementing QoS functionality based on this field.

Is MPLS support for DiffServ possible in cell-mode MPLS knowing that there is no MPLS Shim header and therefore no EXP bits are available?

A. In this case, the DSCP in the IP packet gets encoded in the MPLS cell using the label (VPI/VCI) and the CLP bit in the ATM header. The portion of the DSCP that defines the packet class is encoded in the MPLS label (VPI/VCI in this case.) The portion of the DSCP that defines a drop precedence (when it exists) is encoded in the CLP bit. LSPs using this approach are a special case of L-LSPs.

What is an E-LSP?

A. An E-LSP is an LSP on which nodes infer the QoS treatment for MPLS packet exclusively from the EXP bits in the MPLS header. Because the QoS treatment is inferred from the EXP (both class and drop precedence), several classes of traffic can be multiplexed onto a single LSP (use the same label). A single LSP can support up to eight classes of traffic because the EXP field is a 3-bit field. The maximum number of classes would be less after reserving some values for control plane traffic or if some of the classes have a drop precedence associated with them. E-LSPs are not an option for ATM-LSR. On those devices, MPLS packets use the original ATM cell encapsulation and no Experimental field exists in the cell header.

What is an L-LSP?

A. An L-LSP is an LSP on which nodes infer the QoS treatment for MPLS packets from the packet label and the EXP bits (or the CLP bit for cell-mode MPLS.) In particular, the label is used to encode the class a packet belongs to and the EXP field (or the CLP bit for cell-mode MPLS) is used to encode the drop precedence of the packet. A separate LSP can be established for each combination of FEC and class. For example, three separate LSPs will be established to a single destination if there



are packets belonging to three different classes that reach that destination. The class associated with an L-LSP needs to be signaled explicitly during label establishment so each LSR can subsequently infer the packet class from the label. A new RSVP object (DIFFSERV) and a new LDP TLV (DiffServ) are defined for this purpose.

Are L-LSPs better than E-LSPs? Do E-LSPs provide a very limited solution with just eight possible markings (classes)?

A. The best alternative depends on the details of the QoS design. In particular, it depends on the number of classes, drop precedence values and whether ATM-LSRs are used or not. For ATM-LSR, L-LSPs are the only option. For MPLS deployments that don't use ATM LSR, it's the number of classes and drop precedence values that would determine whether E-LSPs or L-LSP are required. However, service providers that are currently deploying QoS are using at most 4 classes. The three bits in the EXP bits provide enough values for encoding 4 classes (some of them with a drop precedence) and reserving some values for control plane traffic. In this case, E-LSPs satisfy the class requirements. Only deployments with greater class requirements or drop precedence values would need L-LSPs for LSRs other than ATM-LSRs.

How does an E-LSPs or L-LSP differ from a regular LSP?

A. E-LSPs and L-LSP are not inherently different from regular LSPs. They don't define new label binding protocols or modify the packet switching algorithm (label swapping) defined for MPLS packets. They are LSPs on which a particular mechanism of implementing QoS using DiffServ is being used.

How are the experimental bits set?

A. The edge LSR that imposes the MPLS header needs to set the experimental field to some value. By default, Cisco IOS<sup>®</sup> Software copies the three most significant bits of the DiffServ code point (DSCP) or the IP precedence of the IP packet to the EXP field in the MPLS shim header. This action happens when the MPLS header is initially imposed on the IP packet. Another option is to define a mapping between the DSCP or IP precedence and the EXP bits. This mapping can be configured using the `set mpls experimental` or the `police` commands in the Modular QoS CLI (MQC). This sample configuration defines a mapping between DSCP values and EXP bits:

```
!  
class-map match-all PREMIUM  
  match ip dscp ef  
class-map match-all BUSINESS  
  match ip dscp af31 af32 af33  
!  
policy-map IN-POLICY  
  class PREMIUM  
    set mpls experimental 5  
  class BUSINESS  
    set mpls experimental 4  
  class class-default  
    set mpls experimental 0  
!  
interface Serial0/0  
  ip address 10.32.14.2 255.255.255.0  
  service-policy input IN-POLICY
```

Here is another sample configuration for a mapping between DSCP values and EXP bits using the `police` command:

```
!  
class-map match-all PREMIUM  
  match ip dscp ef  
class-map match-all BUSINESS  
  match ip dscp af31 af32 af33  
!  
policy-map IN-POLICY
```

```

class PREMIUM
  police 128000 4000 4000 conform-action set-mpls-exp-transmit 5 exceed-action drop
class BUSINESS
  police 256000 8000 8000 conform-action set-mpls-exp-transmit 4 exceed-action
  set-mpls-exp-transmit 3
class class-default
  set mpls experimental 0
!
interface Serial0/0
  ip address 10.32.14.2 255.255.255.0
  service-policy input IN-POLICY

```

How can a proper mapping between DSCP (or IP precedence) and the experimental field be implemented?

A. The definition of a proper mapping really depends on the end-to-end QoS design and what one is trying to accomplish. There is not a standard that defines the semantics associated with each of the EXP field values. Cisco generally recommends you not use values 6 and 7 for traffic other than control plane traffic (such as routing protocols). Cisco also suggests that decreasing values are associated with decreasing priorities.

This is an example of a hypothetical mapping between DSCP and EXP bits in which four classes of service are defined: one for real-time traffic, another for best-effort traffic, and two more metered classes that incorporate the concept of in-profile and out-of-profile traffic.

Class	DSCP	EXP
Reserved for control plane traffic	Class Selector 7	7
Reserved for control plane traffic	Class Selector 6	6
Class 1 (real-time traffic)	EF	5
Class 2 in-profile	AF31	4
Class 2 out-of-profile	AF32, AF33	3
Class 3 in-profile	AF11	2
Class 3 out-of-profile	AF12, AF13	1
Class 4 (best effort)	Default	0

Here is another example in which six classes of traffic are defined and a one-to-one mapping is implemented between IP precedence values and EXP bits.

Class	IP Precedence	EXP
Reserved for control plane traffic	7	7
Reserved for control plane traffic	6	6
Class 1 (real-time traffic)	5	5
Class 2	4	4
Class 3	3	3
Class 4	2	2
Class 5	1	1
Class 6 (best effort)	0	0



In this third example, four classes of service are defined. Class 1 is used for real-time traffic, class 4 is used for best-effort traffic, and out of the two additional classes, one is a metered class that defines three levels of drop precedence equivalent to the three levels of drop precedence defined for AF in DiffServ.

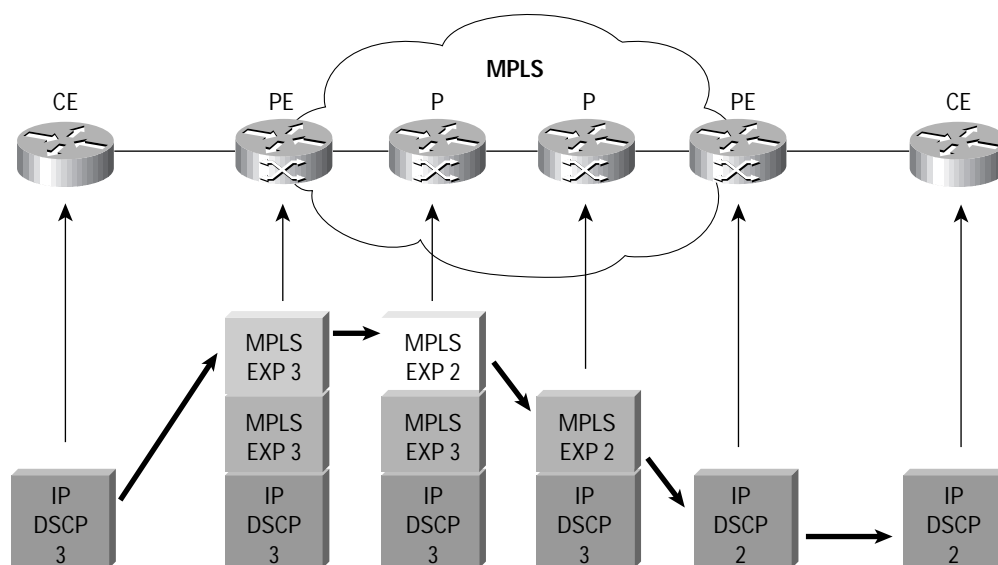
Class	DSCP	EXP
Reserved for control plane traffic	Class Selector 7	7
Reserved for control plane traffic	Class Selector 6	6
Class 1 (real-time traffic)	EF	5
Class 2	Class Selector 4	4
Class 3 (conforming traffic)	AF31	3
Class 3 (exceeding traffic)	AF32	2
Class 3 (violating traffic)	AF33	1
Class 4 (best effort)	Default	0

What are the MPLS QoS tunnel modes?

A. The specification for MPLS support for DiffServ defines three different models of operation with specific rules of interaction between the DiffServ marking in the IP header and the DiffServ marking in the MPLS header. These three tunnel modes are called: Uniform, Pipe, and Short Pipe. Using these tunnel modes, an MPLS network can tunnel transparently the DiffServ marking of packets while defining a separate DiffServ domain or the MPLS network can act as part of a greater DiffServ domain. These tunnel modes only affect the behavior of edge and penultimate LSRs where labels are pushed and popped. They do not affect the label swapping process at intermediate nodes.

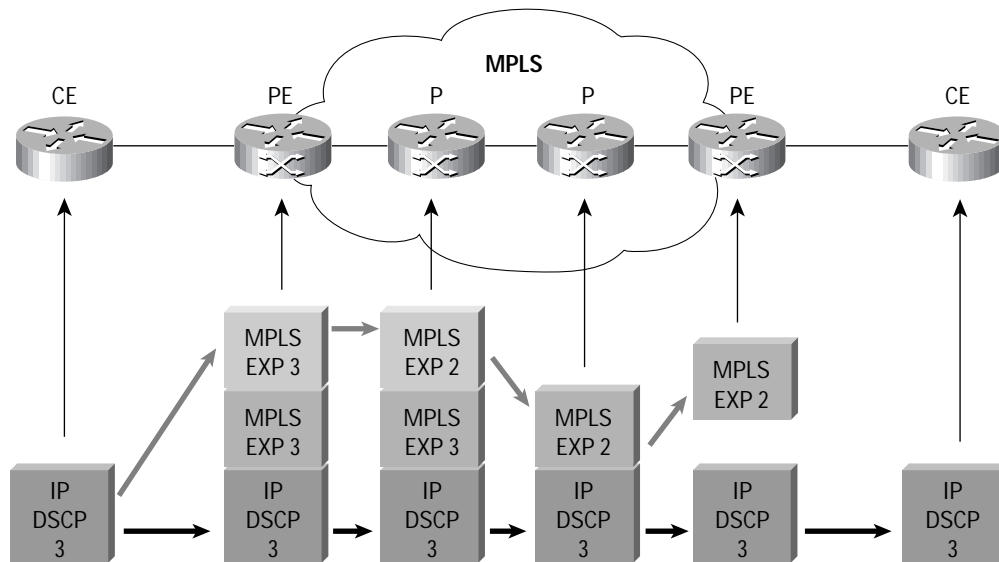
How does the Uniform tunnel mode work?

A. In this tunnel mode, there is only one DiffServ marking that is relevant for a packet when traversing the MPLS network. If the DiffServ marking of the packet is modified within the MPLS network (such as due to traffic conditioning), the updated information is the one considered meaningful at the egress of the LSP. Any changes to the packet marking within the MPLS network are permanent and get propagated when the packet leaves the MPLS network.



How does the Pipe tunnel mode work?

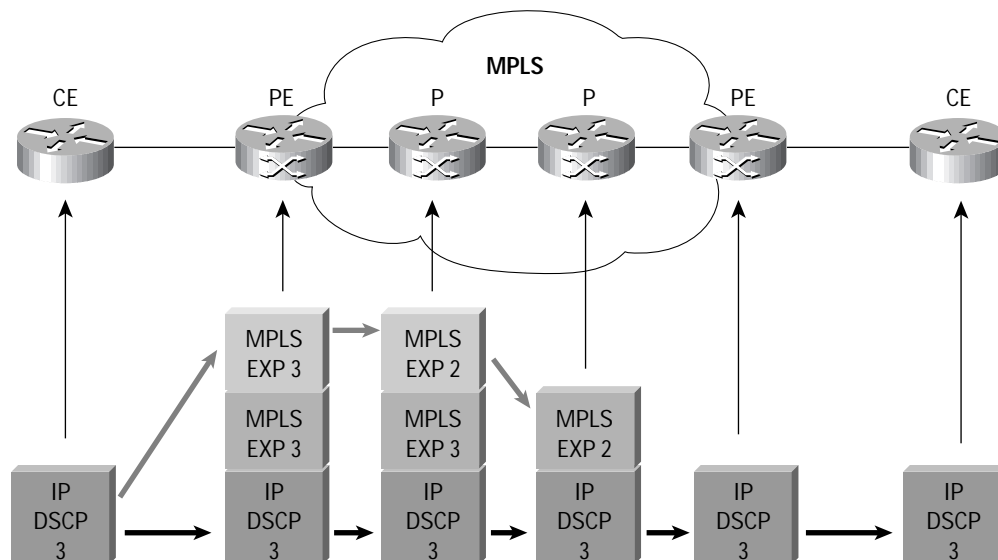
A. In this tunnel mode, two markings are relevant for a packet when traversing the MPLS network. First, the marking used by intermediate nodes along the LSP span including the egress LSR. Second, the original marking carried by the packet before entering the MPLS network that will continue to be used once the packet leaves the MPLS network. Any changes to the packet marking within the MPLS network are not permanent and do not get propagated when the packet leaves the MPLS network. Note that the egress LSR still uses the marking that was used by intermediate LSRs. However, the egress LSR has to remove all labels imposed on the original packet. In order to preserve this marking carried in the labels, the edge LSR keeps an internal copy of the marking before removing the labels. This internal copy is used to classify the packet on the outbound interface (facing the CE) once the labels are removed.





How does the Short-Pipe tunnel mode work?

A. The Short-Pipe tunnel mode is a slight variation of the Pipe tunnel mode. The only difference is that the egress LSR uses the original packet marking instead of using the marking used by the intermediate LSRs.



Can MPLS QoS and MPLS VPNs be used simultaneously? Can they interact?

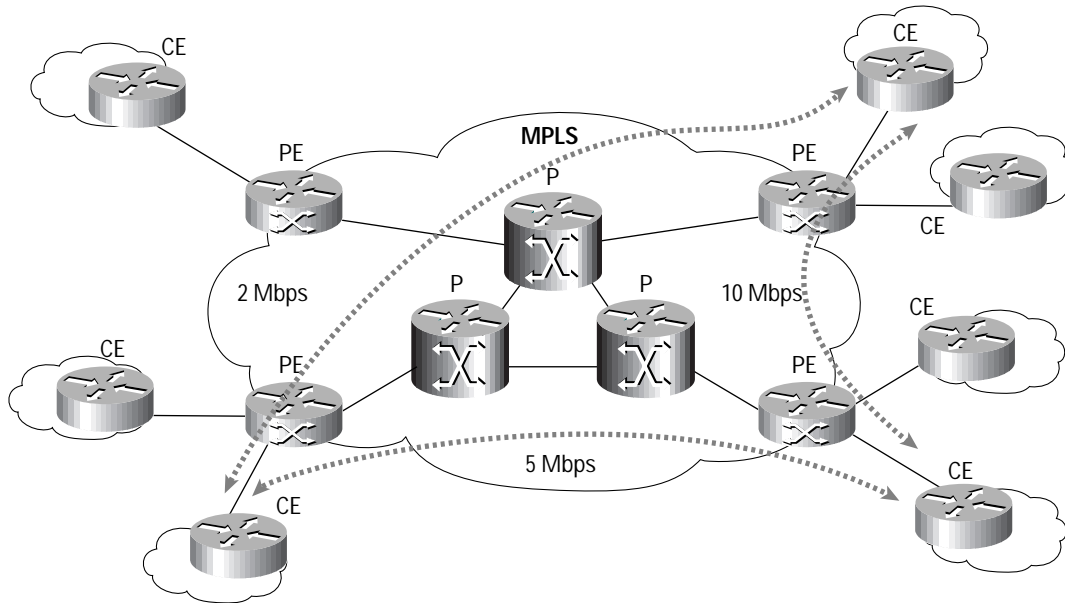
A. MPLS QoS is implemented using the DiffServ architecture regardless of other MPLS applications. Furthermore, MPLS QoS acts as a base technology that enhances other MPLS applications (such as a VPN, TE, and so on).

Can you implement per-VPN QoS in an MPLS network?

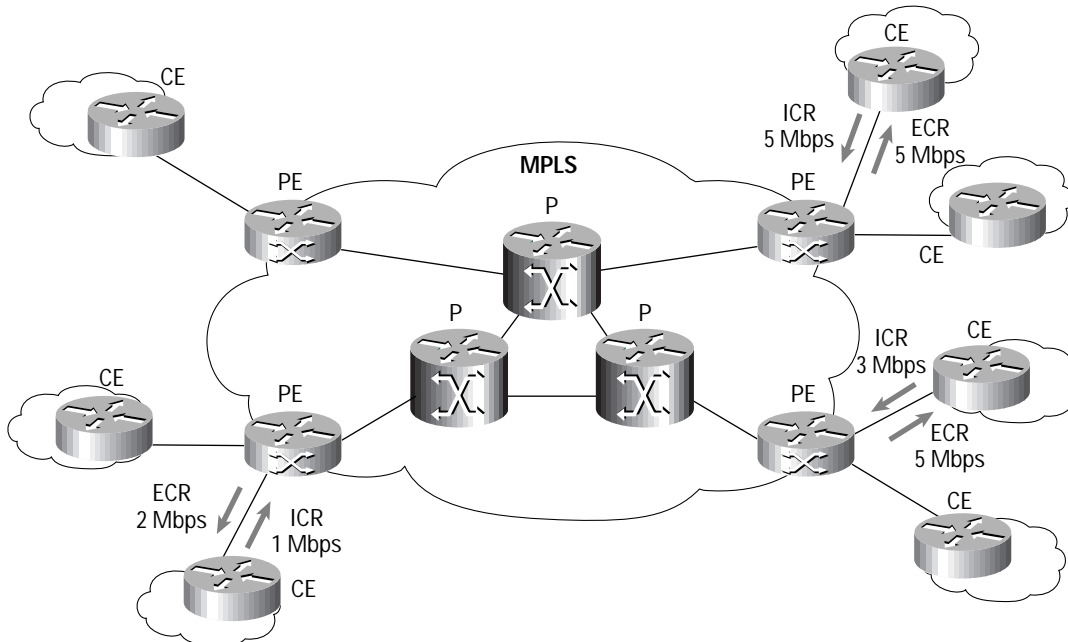
A. Per-VPN QoS is achieved by implementing different ingress QoS policies for different VPNs on edge LSRs. In an MPLS network, all the VPN knowledge resides on those devices. After traffic is switched into the MPLS network, no per-VPN differentiation is performed. Therefore, no per-VPN queuing or dropping of traffic is implemented in the core of the network. The same QoS policies are applied to all packets regardless of the VPN they belong to. With this approach, a highly scalable QoS solution is possible regardless of the number of VPNs.

What are the point-to-cloud and point-to-point QoS models?

A. These are models used in the context of MPLS VPNs that refer to how QoS guarantees are offered. In the point-to-point model, VPN sites have specific QoS guarantees to other VPN sites. In a VPN with three sites, each site will have specific QoS guarantees to the other two sites that belong to the VPN. This approach can be used to offer hard QoS guarantees equivalent to the guarantees available with ATM or Frame Relay.



In the point-to-cloud model, each site receives a single QoS guarantee for traffic sent to and received from all other VPN sites. Two parameters are defined for this purpose, Ingress Committed Rate (ICR) and Egress Committed Rate (ECR). In a VPN with three sites, each site will have a single QoS guarantee for all incoming traffic (regardless of the traffic source) and all outgoing traffic (regardless of the traffic destination).





These two models are not mutually exclusive. A service provider can offer VPN services to a customer where some sites have a hard (point-to-point) guarantee to some other sites while yet other sites have just a soft (point-to-cloud) guarantee. You may find the point-to-cloud model being called the hose model and the point-to-point model being called the Pipe model.

How are those models related to Uniform, Pipe, and Short-Pipe tunnel modes?

A. They are two different classifications. The Uniform, Pipe, and Short-Pipe tunnel modes define the rules of interaction between the original DiffServ marking of an IP packet and the marking used within the MPLS network. These rules are independent of the MPLS application used. The point-to-cloud and point-to-point QoS models are only relevant in the context of MPLS VPNs and refer to how the QoS guarantees between sites are implemented.

Can MPLS QoS and MPLS TE be used simultaneously? Can they interact?

A. MPLS QoS is implemented using the DiffServ architecture regardless of other MPLS applications. Furthermore, MPLS QoS acts as a base technology that enhances other MPLS applications (such as a VPN, TE, and so on).

What is DiffServ-Aware Traffic Engineering (DS-TE)?

A. DS-TE is an enhancement to MPLS TE that introduces the concept of classes (class types to be more precise) to TE. Each participating link advertises the amount of available bandwidth of each class type on that link. When the constrain-based-routing process is executed for a new tunnel, a bandwidth constrain of a particular class type can be defined as one of the criteria to be used for the path selection. The admission control process carried using RSVP at each hop is performed against the available bandwidth of the specific class type. For example, you can define a tunnel between point A and B of the network with 10 Mbps of available bandwidth of a particular class type. The current Cisco IOS Software implementation supports two class types. These two bandwidth amounts are not treated as two separate amounts. Instead, each link defines a bandwidth pool and bandwidth subpool. The bandwidth subpool is part of the bandwidth pool. As a nonoversubscribed example, an OC-3/STM-1 can have a bandwidth pool with 155 Mbps and a bandwidth subpool with 55 Mbps.

Are DS-TE and L-LSP similar mechanisms to provide separate LSPs for different classes of service between two points?

A. DS-TE and L-LSP are both mechanisms that can provide different LSPs for different classes of service, but they have different characteristics and serve different purposes. For example, L-LSPs could be used on an MPLS network that does not implement TE. On the other hand, DS-TE offers constrain-based routing capabilities that are DiffServ aware. This functionality cannot be provided with L-LSPs. However, they are not mutually exclusive. Potentially, a DS-TE network could use L-LSP.

Why can hard QoS guarantees be implemented with DS-TE?

A. DS-TE offers the control plane infrastructure that coupled with the DiffServ forwarding plane features make hard QoS guarantees possible. Using DS-TE, you can define a path between two points that has a bandwidth guarantee of a particular class type that satisfies certain QoS requirements (delay, jitter, and loss). Those QoS requirements are satisfied along the LSP span using queuing and dropping policies. With the correct provisioning, the admission control process in DS-TE ensures that at each hop the queuing and dropping mechanisms in the forwarding plane are not over committed and the QoS guarantees (delay/jitter/packet loss) and bandwidth guarantees are met.

What kind of applications does DS-TE enable?

A. The most obvious application for DS-TE is the implementation of voice trunks across an MPLS network. DS-TE can satisfy the strict service requirements of voice traffic by providing strong QoS guarantees for bandwidth, delay, jitter, and packet loss. Using DS-TE, a service provider can offer voice over MPLS (VoMPLS) services. This service transports voice as VoIP traffic that is carried inside the MPLS network. DS-TE could also be used for other premium services trunking traffic with specific QoS and bandwidth guarantees across an MPLS backbone.

Why does Penultimate-Hop Popping (PHP) affect the implementation of MPLS QoS?

A. In some cases, the PHP can expose the IP packet to the penultimate hop. This happens when an MPLS packet arriving at the penultimate hop has only one label. In this case, the penultimate LSR and the edge LSR do not have access to the EXP value that the packet carried before the MPLS header was removed. To preserve the EXP value in this case, the edge LSR needs to advertise an explicit NULL label (a label value of zero). The penultimate hop forwards MPLS packets with a NULL label instead of forwarding IP packets. An explicit NULL label is not needed when the penultimate hop receives MPLS packets with a label stack that contains at least two labels and PHP is performed. In that case, the inner label can still carry the EXP value needed by the penultimate and edge LSR to implement their QoS policy.

Where can I find additional information on MPLS QoS?

Cisco IOS MPLS

<http://www.cisco.com/go/mpls>

Cisco IOS MPLS Configuration Guide, Release 12.2

[http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fswtch\\_c/swprt3/index.htm](http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fswtch_c/swprt3/index.htm)

Cisco IOS Switching Services Command Reference, Release 12.2

[http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fswtch\\_r/index.htm](http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fswtch_r/index.htm)

Cisco IOS Quality of Service Solutions Configuration Guide, Release 12.2

[http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos\\_c/index.htm](http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/index.htm)

Cisco IOS Quality of Service Solutions Command Reference, Release 12.2

[http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos\\_r/index.htm](http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_r/index.htm)

MPLS: Technology and Applications. Bruce Davie and Yakov Rekhter. ISBN 1-55860-656-4



Corporate Headquarters  
Cisco Systems, Inc.  
170 West Tasman Drive  
San Jose, CA 95134-1706  
USA  
[www.cisco.com](http://www.cisco.com)  
Tel: 408 526-4000  
800 553-NETS (6387)  
Fax: 408 526-4100

European Headquarters  
Cisco Systems Europe  
11, Rue Camille Desmoulins  
92782 Issy-les-Moulineaux  
Cedex 9  
France  
[www-europe.cisco.com](http://www-europe.cisco.com)  
Tel: 33 1 58 04 60 00  
Fax: 33 1 58 04 61 00

Americas Headquarters  
Cisco Systems, Inc.  
170 West Tasman Drive  
San Jose, CA 95134-1706  
USA  
[www.cisco.com](http://www.cisco.com)  
Tel: 408 526-7660  
Fax: 408 527-0883

Asia Pacific Headquarters  
Cisco Systems Australia, Pty., Ltd  
Level 9, 80 Pacific Highway  
P.O. Box 469  
North Sydney  
NSW 2060 Australia  
[www.cisco.com](http://www.cisco.com)  
Tel: +61 2 8448 7100  
Fax: +61 2 9957 4350

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the

**Cisco Web site at [www.cisco.com/go/offices](http://www.cisco.com/go/offices)**

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia  
Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia  
Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru  
Philippines • Poland • Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa  
Spain • Sweden • Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Copyright © 2001, Cisco Systems, Inc. All rights reserved. Cisco, Cisco IOS, Cisco Systems, and the Cisco Systems logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0108R)

LW2696 10/01

Printed in the USA