CHAPTER 5

MPLS VPN QoS Design

MPLS VPNs are rapidly gaining popularity as private-WAN alternatives. The migration to a MPLS VPN from a private-WAN requires a significant paradigm shift when addressing QoS designs. This is because enterprise customer subscribers must closely cooperate with their service providers to ensure end-to-end service-levels; they can no longer achieve these service-levels independent of their service provider's policies.

MPLS VPN QoS design can be viewed from two distinct perspectives:

- The enterprise customer subscribing to the MPLS VPN service
- The service provider provisioning edge and core QoS within the MPLS VPN service

To achieve end-to-end service levels, both enterprise and service-provider QoS designs must be consistent and complimentary. This design chapter will focus primarily on the enterprise customer's perspective, yet elements of the service provider's side of the equation will also be included to round out the picture and to better convey how the two sides fit together. Furthermore, as some enterprises self-manage their MPLS VPNs, it is important to examine both elements of the QoS solution.

The following topics are discussed in this design chapter:

- Enterprise-to-Service Provider Mapping Models
- Service Provider-to-Enterprise Models
- MPLS DiffServ Tunneling Modes

MPLS is a combination of routing and switching technologies that can provide scalable VPNs with end-to-end quality of service.

Many enterprise customers are turning to service providers that offer MPLS VPN services as private WAN alternatives. One of the main reasons for this is the any-to-any connectivity capabilities of MPLS VPNs. However, this full-mesh nature in itself poses significant QoS implications to enterprise customers and service providers alike—namely, that they both need to comanage QoS in a cooperative and complementary fashion to achieve end-to-end service levels.

This chapter examines in detail QoS considerations that enterprise customers need to bear in mind when subscribing to MPLS VPNs, including how best to map into various service-provider MPLS VPN QoS models.

Service provider-edge QoS considerations are also presented, including egress queuing models and MPLS DiffServ tunneling modes (Uniform, Short Pipe, and Pipe).

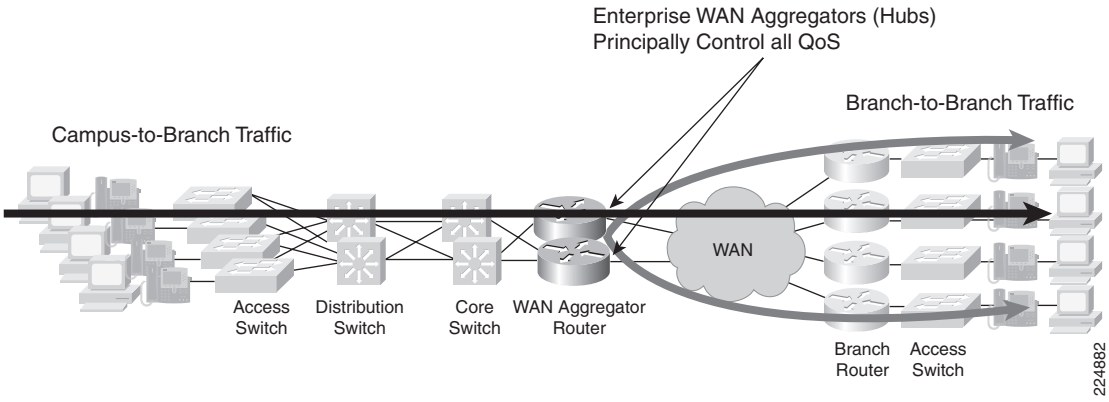
Note This chapter addresses QoS design for MPLS VPNs, not the theory and operation of MPLS VPNs themselves. It is assumed that the reader is familiar with basic MPLS VPN architectures and technologies. For a detailed discussion of MPLS VPNs, refer to the Cisco Press books *MPLS and VPN Architectures*, Volumes I and II, by Ivan Pepelnjak and Jim Guichard; *Traffic Engineering with MPLS*, by Eric Osborne and Ajay Simha; and *Advanced MPLS Design and Implementation*, by Vivek Alwayn.

Where Is QoS Needed over an MPLS VPN?

MPLS VPN architectures are comprised of customer edge (CE) routers, provider-edge (PE) routers, and provider (P) routers. MPLS VPNs provide fully meshed Layer 3 virtual WAN services to all interconnected CE routers, as outlined by RFC 2547. This fully meshed characteristic of MPLS VPNs presents a significant design implication to traditional Layer 2 WAN QoS design.

Because of cost, scalability, and manageability constraints, traditional private WAN designs rarely use full-mesh models. Instead, most Layer 2 WAN designs revolve around a hub-and-spoke model, implementing either a centralized hub design or the more efficient regional hub design. Under such hub-and-spoke designs, QoS primarily is administered at the hub router by the enterprise. As long as the service provider meets the contracted service levels, the packets received at remote branches will reflect the scheduling policies of the hub router (sometimes referred to as a *WAN aggregator*). The WAN aggregator controls not only campus-to-branch traffic, but also branch-to-branch traffic (which is homed through the hub). Under traditional hub-and-spoke models, QoS principally is administered by the enterprise customer, as shown in Figure 5-1.

Figure 5-1 QoS Administration in Traditional Hub-and-Spoke Layer 2 WAN Design



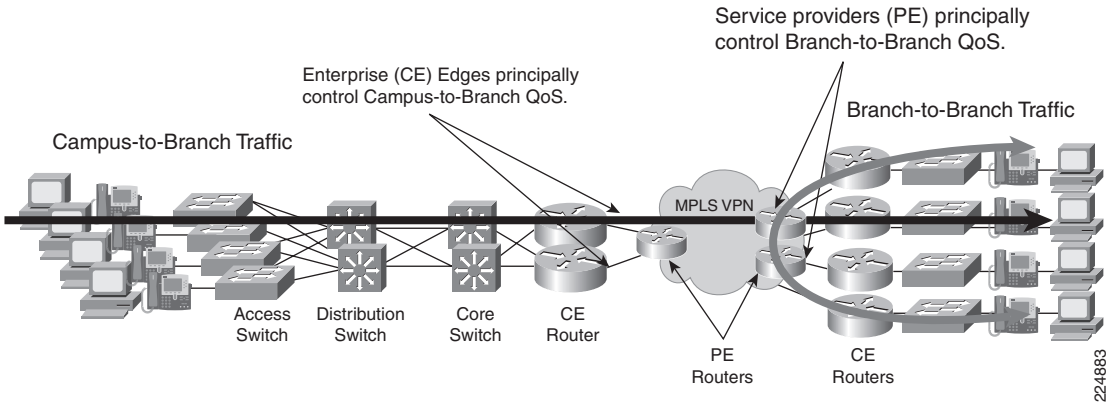
However, with the advent of MPLS VPN service offerings that inherently offer full-mesh connectivity, the QoS administration paradigm shifts. Under a full-mesh design, the hub router still administers QoS for all campus-to-branch traffic, but it no longer fully controls the QoS for branch-to-branch traffic. Although it might appear that the only required workaround for this new scenario is to ensure that QoS is provisioned on all branch routers, this is insufficient because it addresses only part of the issue.

For example, consider the case of provisioning any-to-any videoconferencing. As with a traditional Layer 2 WAN design, a scheduling policy to prioritize IP/VC on the WAN aggregator is required. Then the enterprise must properly provision similar priority scheduling for IP/VC on the branch routers also. In this manner, any videoconferencing calls from the campus to the branch (and also from branch to branch) are protected against traffic of lesser importance flowing between the *same* sites. The complexity of the fully meshed model arises when considering that contending traffic might not always

come for the same sites, but could come from *any* site. Furthermore, the enterprise no longer fully controls QoS for branch-to-branch traffic because this traffic no longer is homed through a hub. Continuing the example, if a videoconferencing call is set up between two branches and a user from one of the branches also initiates a large FTP download from the central site, the potential for oversubscription of the PE-to-CE link from the fully meshed MPLS VPN cloud into one of the branches becomes very real, likely causing drops from the IP/VC call.

The only way to guarantee service levels in such a scenario is for the service provider to provision QoS scheduling that is compatible with the enterprise's policies on all PE links to remote branches. This is what creates the paradigm shift in QoS administration for fully meshed topologies. Namely, enterprises and service providers must cooperate to jointly administer QoS over MPLS VPNs, as shown in Figure 5-2.

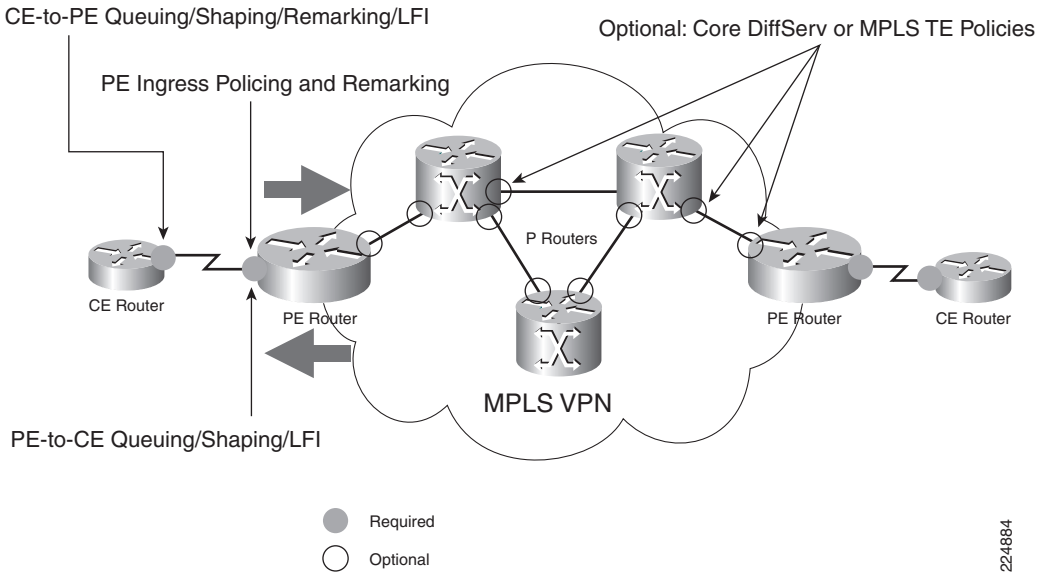
Figure 5-2 QoS Administration in Fully Meshed MPLS VPN Design



Queuing policies are mandatory on CE and PE routers because of the full-mesh implications of MPLS VPNs. PE routers also typically have policing (and markdown) policies on ingress to enforce SLAs.

QoS policies on P routers are optional. Such policies are optional because some service providers overprovision their MPLS core networks and, as such, do not require any additional QoS policies within their backbones; on the other hand, other providers might implement simplified DiffServ policies within their cores or might even deploy MPLS traffic engineering (MPLS TE) to handle congestion scenarios within their backbones. Figure 5-3 summarizes the points where QoS policies can be provisioned within MPLS VPN architectures.

Figure 5-3 Where QoS Is Required in MPLS VPN Architectures



This design chapter focuses on CE and PE QoS design.

Customer Edge QoS Design Considerations

In addition to the full-mesh implication of MPLS VPNs, these considerations should be kept in mind when considering MPLS VPN CE QoS design:

- Layer 2 access (link-specific) QoS design
- Service-provider service-level agreements (SLA)
- Enterprise-to-service provider mapping models

The following sections examine these considerations in more detail.

Layer 2 Access (Link-Specific) QoS Design

Although MPLS VPNs are essentially Layer 3 WANs, a Layer 2 access medium to connect to the MPLS VPN service provider is an obvious requirement. Most providers support Frame Relay and ATM as access media because this makes migration from Layer 2 WANs to Layer 3 MPLS VPNs easier and cheaper to manage; customers are not forced to convert hardware on hundreds (or, in some cases, thousands) of remote branch routers to connect to MPLS VPN providers.

It is important to recognize that Layer 2 QoS link-specific issues and designs remain the same with regular Layer 2 WAN edges or with Layer 3 MPLS VPN CE/PE edges. For example, shaping and LFI recommendations for slow-speed FR links are identical whether the link is used for a Layer 2 WAN or for a Layer 3 MPLS VPN access link. Again, this makes migration easier to manage because link-specific QoS designs do not need to be changed (although the service policy itself might require minor modification, which is discussed in more detail shortly).

In addition to FR and ATM for access, some service providers support Ethernet/Fast Ethernet as access media but usually guarantee a CIR of only subline rate. In such cases, hierarchical shaping and queuing policies on the CE edges are recommended, as illustrated later in this chapter.

Service Provider Service-Level Agreements

End-to-end QoS is like a chain that is only as strong as the weakest link. Therefore, it's essential for enterprises (with converged networks) subscribing to MPLS VPN services to choose service providers that can provide the required SLAs for their converged networks. For example, these are the end-to-end SLA requirements of voice and interactive video:

- No more than 150 ms of one-way latency from mouth to ear (per ITU G.114 standard)
- No more than 30 ms of jitter
- No more than 1 percent loss

As a subset of the trip, the service provider's component of the SLA must be considerably tighter. These SLAs are defined for Cisco-Powered Networks (CPN)–IP Multiservice Service Providers:

- No more than 60 ms of one-way latency from edge to edge
- No more than 20 ms of jitter
- No more than 0.5 percent loss

Figure 5-4 illustrates the interrelationship of these SLAs.

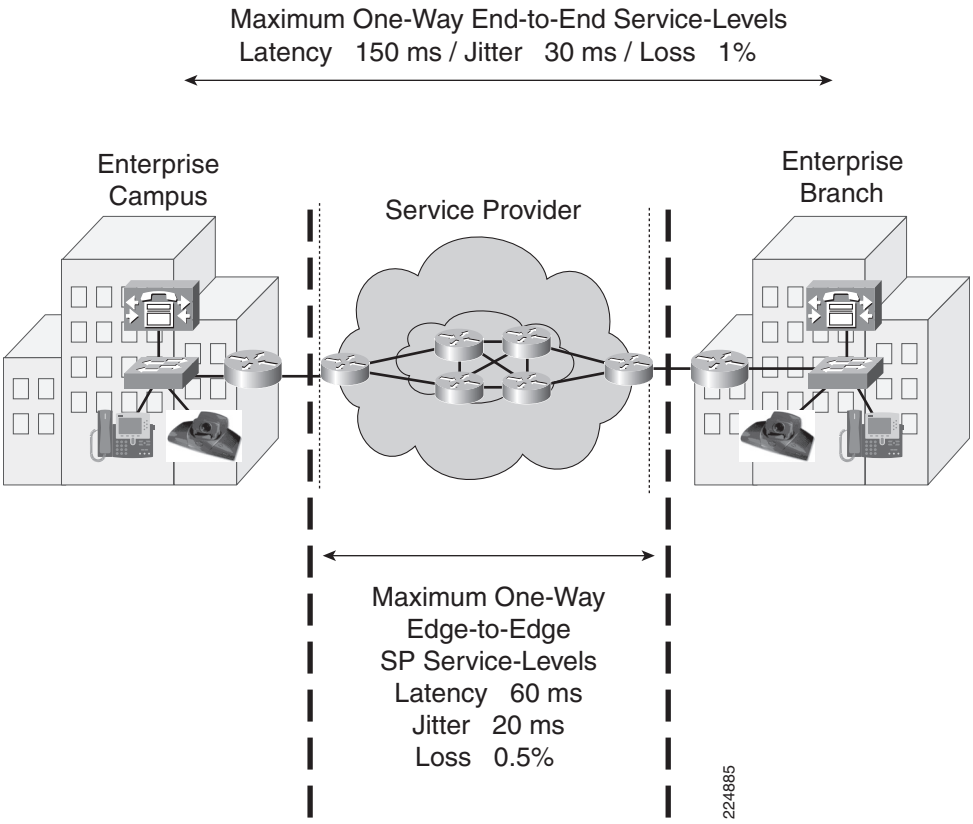
CPN-IP Multiservice Service Providers that meet these SLAs can be found at:

http://www.cisco.com/pcgi-bin/cpn/cpn_pub_bassrch.pl

Choose the IP VPN Multiservice option.

To achieve such end-to-end SLAs, enterprise customers (managing CEs) and service providers (managing PEs and core Ps) must cooperate and be consistent in classifying, provisioning, and integrating their respective QoS designs. To this end, various mapping models have been developed to integrate enterprise requirements into service-provider solutions.

Figure 5-4 CPN-IP Multiservice Service-Provider SLAs



Enterprise-to-Service Provider Mapping Models

Although Cisco is adopting its new QoS Baseline initiative and designing tools such as Cisco AutoQoS Enterprise to facilitate and simplify the deployment of advanced QoS traffic models within the enterprise, to date, very few enterprises have deployed more than a handful of traffic classes. Therefore, most service providers offer only a limited number of classes within their MPLS VPN clouds. At times, this might require enterprises to collapse the number of classes that they have provisioned to integrate into their service provider's QoS models. The following caveats should be remembered when deciding how best to collapse and integrate enterprise classes into various service-provider QoS models.

Voice and Video

Service providers typically offer only one Real-Time class or Priority class of service. If an enterprise wants to deploy both Voice and Interactive-Video (each of which is recommended to be provisioned with strict priority treatment) over their MPLS VPN, they might be faced with a dilemma. Which one should be assigned to the Real-Time class? Are there any implications about assigning both to the Real-Time class?

Keep in mind that voice and video should never both be assigned low-latency queuing on link speeds where serialization is a factor (≤768 kbps). Packets offered to the LLQ typically are not fragmented; thus, large IP/VC packets can cause excessive delays for VoIP packets on slow-speed links.

An alternative is to assign IP/VC to a nonpriority class, which entails not only the obvious caveat of lower service levels, but also possible traffic-mixing concerns, as discussed shortly.

Call-Signaling

VoIP requires provisioning not only of RTP bearer traffic, but also of Call-Signaling traffic, which is very lightweight and requires only a moderate amount of guaranteed bandwidth. Because the service levels applied to Call-Signaling traffic directly affect delay to the dial tone, it is important from the end user's expectations that Call-Signaling be protected. Service providers might not always offer a suitable class just for call signaling traffic itself, leading to the question of which other traffic classes Call-Signaling should be mixed with.

On links where serialization is not an issue (>768 kbps), Call-Signaling could be provisioned into the Real-Time class, along with voice.

However, this is not recommended on slow-speed links where serialization *is* a factor. On such slow-speed links, Call-Signaling is best assigned to one of the preferential data classes for which the service provider provides a bandwidth guarantee.

It is important to realize that a guarantee applied to a service-provider class as a whole does not itself guarantee adequate bandwidth for an individual enterprise applications within the class.

Mixing TCP with UDP

It is a general best practice to not mix TCP-based traffic with UDP-based traffic (especially Streaming-Video) within a single service-provider class because of the behaviors of these protocols during periods of congestion. Specifically, TCP transmitters throttle back flows when drops are detected. Although some UDP applications have application-level windowing, flow control, and retransmission capabilities, most UDP transmitters are completely oblivious to drops and, thus, never lower transmission rates because of dropping.

When TCP flows are combined with UDP flows within a single service-provider class and the class experiences congestion, TCP flows continually lower their transmission rates, potentially giving up their bandwidth to UDP flows that are oblivious to drops. This effect is called TCP starvation/UDP dominance.

TCP starvation/UDP dominance likely occurs if (TCP-based) Mission-Critical Data is assigned to the same service-provider class as (UDP-based) Streaming-Video and the class experiences sustained congestion. Even if WRED is enabled on the service-provider class, the same behavior would be observed because WRED (for the most part) manages congestion only on TCP-based flows.

Granted, it is not always possible to separate TCP-based flows from UDP-based flows, but it is beneficial to be aware of this behavior when making such application-mixing decisions within a single service-provider class.

Marking and Re-Marking

Most service providers use Layer 3 marking attributes (IPP or DSCP) of packets offered to them to determine which service provider class of service the packet should be assigned to. Therefore, enterprises must mark or re-mark their traffic consistent with their service provider's admission criteria to gain the appropriate level of service. Additionally, service providers might re-mark at Layer 3 out-of-contract traffic within their cloud, which might affect enterprises that require consistent end-to-end Layer 3 markings.

A general DiffServ principle is to mark or trust traffic as close to the source as administratively and technically possible. However, certain traffic types might need to be re-marked before handoff to the service provider to gain admission to the correct class. If such re-marking is required, it is recommended that the re-marking be performed at the CE's egress edge, not within the campus. This is because service-provider service offerings likely will evolve or expand over time, and adjusting to such changes will be easier to manage if re-marking is performed only at the CE egress edge.

Additionally, in some cases, multiple types of traffic are required to be marked to the same DiffServ code point value to gain admission to the appropriate queue. For example, on high-speed links, it might be desired to send Voice, Interactive-Video, and Call-Signaling to the service provider's Real-Time class. If this service-provider class admits only DSCP EF and CS5, two of these three applications would be required to share a common code point. The class-based marking configuration in Example 5-1 shows how this can be done (in this example, both Interactive-Video and Call-Signaling are re-marked to share DSCP CS5).

Example 5-1 CE (Egress) Enterprise-to-Service Provider Re-Marking Example

```
!
class-map match-any VOICE
 match ip dscp ef
class-map match-all INTERACTIVE-VIDEO
 match ip dscp af41
class-map match-any CALL-SIGNALING
 match ip dscp af31
 match ip dscp cs3
!
policy-map CE-EGRESS-EDGE
 class VOICE
  priority percent 18
 class INTERACTIVE-VIDEO
  priority percent 15
  set ip dscp cs5            ! Interactive-Video is remarked to CS5
 class CALL-SIGNALING
  priority percent 2         ! Call-Signaling gets LLQ for this scenario
  set ip dscp cs5            ! Call-Signaling is also remarked to CS5
!
!
interface Serial1/0
 service-policy output CE-EGRESS-EDGE
!
```

Verification commands:

- **show policy**
- **show policy interface**

Service providers might re-mark traffic at Layer 3 to indicate whether certain flows are out of contract. Although this is consistent with DiffServ standards, such as RFC 2597, it might present minor difficulties to enterprises that require consistent end-to-end Layer 3 marking (typically, for management or accounting purposes). In such cases, the enterprise can choose to apply re-marking policies as traffic is received back from the service provider's MPLS VPN (on the ingress direction of the enterprise's CE).

Class-based marking can be used again because it supports not only access lists for classification, but also Network-Based Application Recognition (NBAR).

Continuing and expanding on the previous example, the enterprise wants to restore the original markings that it set for Interactive-Video and Call-Signaling. Additionally, it wants to restore original markings for Oracle traffic (which it originally marked DSCP 25 and is using TCP port 9000 with) and DLSw+ traffic (originally marked AF21). Both of these data applications were handed off to the service provider

marked as AF21, but they might have been marked down to AF22 within the service-provider cloud. Example 5-2 shows a configuration that enables such re-marking from the MPLS VPN. The "match-all" criteria of the class maps performs a logical AND operation against the potential markings and re-markings, and the access list (or NBAR-supported protocol) that sifts the applications apart. The policy is applied on the same CE link, but in the ingress direction.

Example 5-2 CE (Ingress) Service Provider-to-Enterprise Re-Marking Example

```
!
class-map match-all REMARKED-INTERACTIVE-VIDEO
  match ip dscp cs5
  match access-group 101              ! Interactive-Video must be CS5 AND UDP
!
class-map match-all REMARKED-CALL-SIGNALING
  match ip dscp cs5
  match access-group 102              ! Call-Signaling must be CS5 AND TCP
!
class-map match-all REMARKED-ORACLE
 match ip dscp af21  af22             ! Oracle may have been remarked to AF22
 match access-group 103               ! Oracle uses TCP port 9000
!
class-map match-all REMARKED-DLSW+
 match ip dscp af21  af22             ! DLSw+ may have been remarked to AF22
 match protocol dlsw                  ! DLSw+ is identified by NBAR
!
policy-map CE-INGRESS-EDGE
 class REMARKED-INTERACTIVE-VIDEO
  set ip dscp af41             ! Restores Interactive-Video marking to AF41
 class REMARKED-CALL-SIGNALING
  set ip dscp af31             ! Restores Call-Signaling marking to AF31
 class REMARKED-ORACLE
  set ip dscp 25               ! Restores Oracle marking to DSCP 25
 class REMARKED-DLSW+
  set ip dscp af21             ! Restores DLSw+ marking to AF21
!
!
interface serial 1/0
 service-policy output CE-EGRESS-EDGE
 service-policy input  CE-INGRESS-EDGE       ! Marking restoration on ingress
!
!
access-list 101 permit udp any any           ! Identifies UDP traffic
access-list 102 permit tcp any any           ! Identifies TCP traffic
access-list 103 permit tcp any eq 9000 any   ! Identifies Oracle on TCP 9000
!
```

Verification commands:

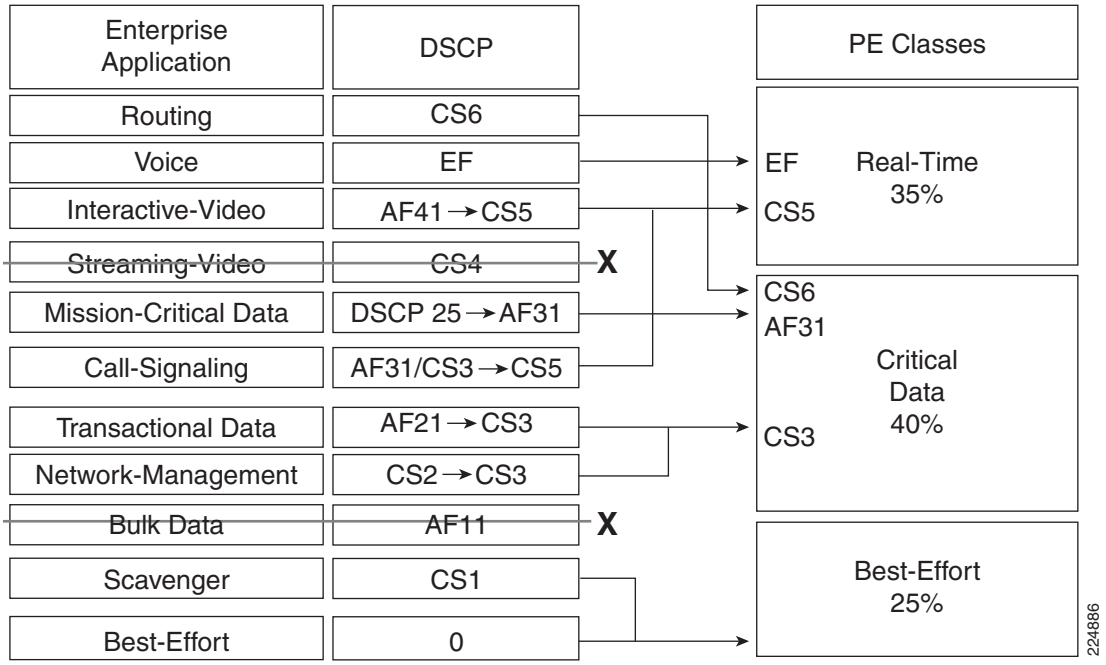
- **show policy**
- **show policy interface**

Three-Class Provider-Edge Model: CE Design

In this model, the service provider offers three classes of service: Real-Time (strict priority, available in 5-percent increments), Critical Data (guaranteed bandwidth), and Best-Effort. The admission criterion for the Real-Time class is either DSCP EF or CS5; the admission criterion for Critical Data is DSCP CS6, AF31, or CS3. All other code points are re-marked to 0. Additionally, out-of-contract AF31 traffic can be marked down within the service provider's MPLS VPN cloud to AF32.

Under such a model, there is no recommended provision for protecting Streaming-Video (following the "Don't mix TCP with UDP" guideline), nor is there a service-provider class suitable for bulk data, which consists of large, nonbursty TCP sessions that could drown out smaller data transactions. Figure 5-5 shows a re-marking diagram for a three-class service-provider model.

Figure 5-5 Three-Class Provider-Edge Model Re-Marking Diagram



Example 5-3 shows an example CE configuration for an advanced enterprise model mapping (over a dual-T1 link) into a three-class service-provider model.

Example 5-3 CE Configuration for Three-Class Provider-Edge Model

```
!
class-map match-all ROUTING
 match ip dscp cs6
class-map match-all VOICE
 match ip dscp ef

class-map match-all INTERACTIVE-VIDEO
 match ip dscp af41
class-map match-all MISSION-CRITICAL-DATA
 match ip dscp 25
class-map match-any CALL-SIGNALING
 match ip dscp af31
 match ip dscp cs3
class-map match-all TRANSACTIONAL-DATA
 match ip dscp af21
class-map match-all NETWORK-MANAGEMENT
 match ip dscp cs2
class-map match-all SCAVENGER
 match ip dscp cs1
!
!
policy-map CE-THREE-CLASS-SP-MODEL
 class ROUTING
```

```
 bandwidth percent 3  ! Routing is assigned (by default) to Critical SP class
class VOICE
 priority percent 18  ! Voice is admitted to Realtime SP class
class INTERACTIVE-VIDEO
 priority percent 15
 set ip dscp cs5      ! Interactive-Video is assigned to the Realtime SP class
class CALL-SIGNALING
 priority percent 2   ! Call-Signaling gets LLQ for this scenario
 set ip dscp cs5      ! Call-Signaling is assigned to the Realtime SP class
class MISSION-CRITICAL-DATA
 bandwidth percent 20
 random-detect
 set ip dscp af31     ! MC Data is assigned to the Critical SP class
class TRANSACTIONAL-DATA
 bandwidth percent 15
 random-detect
 set ip dscp cs3      ! Transactional Data is assigned to Critical SP class
class NETWORK-MANAGEMENT
 bandwidth percent 2
 set ip dscp cs3       ! Net Mgmt is assigned to Critical SP class
class SCAVENGER
 bandwidth percent 1
class class-default
 bandwidth percent 24
 random-detect
!
```

Verification commands:

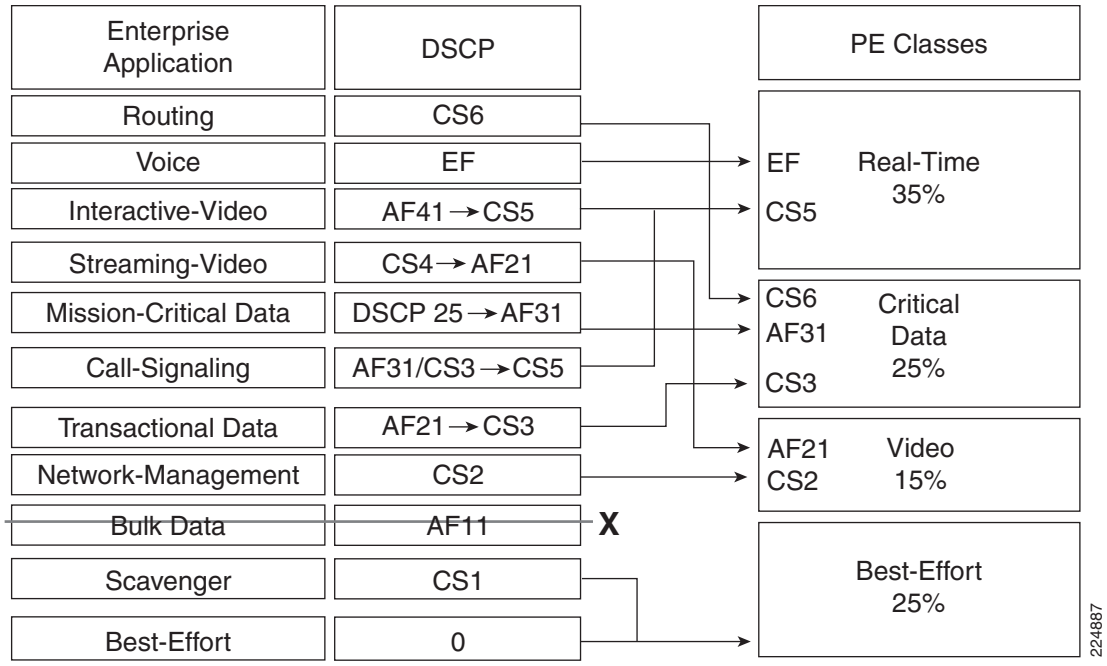
- **show policy**
- **show policy interface**

The **max-reserved-bandwidth** command might be required on the interface to which the previously discussed policy is applied.

Four-Class Provider-Edge Model: CE Design

Building on the previous model, a fourth class is added that can be used for either Bulk Data or Streaming-Video. The admission criterion for this new class is either DSCP AF21 or CS2. The re-marking diagram shown in Figure 5-6 illustrates how this new class can be used for Streaming-Video and (primarily UDP-based) Network-Management traffic.

Figure 5-6 Four-Class Provider-Edge Model Re-Marking Diagram



Example 5-4 shows an example CE configuration for an advanced enterprise model mapping (over a dual-T1 link) into a four-class service-provider model.

Example 5-4 CE Configuration for Four-Class Provider-Edge Model

```
!
class-map match-all ROUTING
 match ip dscp cs6
class-map match-all VOICE
 match ip dscp ef
class-map match-all INTERACTIVE-VIDEO
 match ip dscp af41
class-map match-all STREAMING-VIDEO
 match ip dscp cs4
class-map match-all MISSION-CRITICAL-DATA
 match ip dscp 25
class-map match-any CALL-SIGNALING
 match ip dscp af31
 match ip dscp cs3
class-map match-all TRANSACTIONAL-DATA
 match ip dscp af21
class-map match-all NETWORK-MANAGEMENT
 match ip dscp cs2
class-map match-all SCAVENGER
 match ip dscp cs1
!
!
policy-map CE-FOUR-CLASS-SP-MODEL
 class ROUTING
  bandwidth percent 3  ! Routing is assigned (by default) to Critical SP class
 class VOICE
  priority percent 18  ! Voice is admitted to Realtime SP class
 class INTERACTIVE-VIDEO
  priority percent 15
  set ip dscp cs5      ! Interactive-Video is assigned to the Realtime SP class
```

```
 class STREAMING-VIDEO
  bandwidth percent 13
  set ip dscp af21      ! Streaming-Video is assigned to the Video SP class
 class CALL-SIGNALING
  priority percent 2    ! Call-Signaling gets LLQ for this scenario
  set ip dscp cs5       ! Call-Signaling is assigned to the Realtime SP class
 class MISSION-CRITICAL-DATA
  bandwidth percent 12
  random-detect
  set ip dscp af31      ! MC Data is assigned to the Critical SP class
 class TRANSACTIONAL-DATA
  bandwidth percent 10
  random-detect
  set ip dscp cs3       ! Transactional Data is assigned to Critical SP class
 class NETWORK-MANAGEMENT
  bandwidth percent 2   ! Net Mgmt (mainly UDP) is admitted to Video SP class
 class SCAVENGER
  bandwidth percent 1
 class class-default
  bandwidth percent 24
  random-detect
!
```

Verification commands:

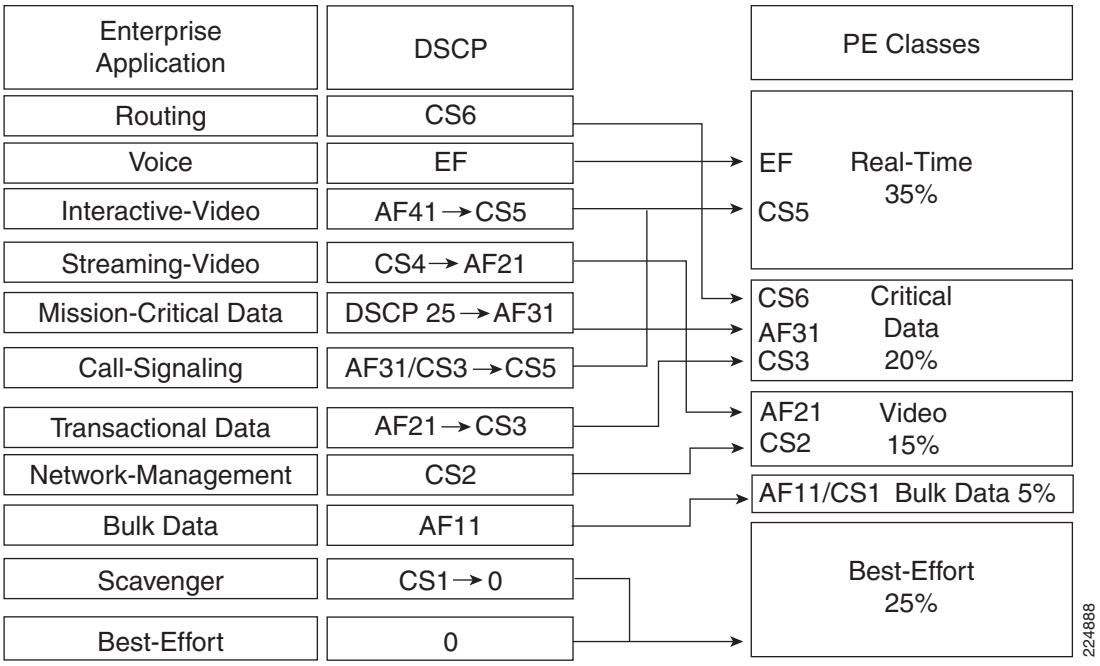
- **show policy**
- **show policy interface**

The **max-reserved-bandwidth** command might be required on the interface to which the previously discussed policy is applied.

Five-Class Provider-Edge Model: CE Design

Building again on the previous model, a fifth class is added that also can be used for either Bulk Data or Streaming-Video (whichever wasn't used under the four-class model). The admission criterion for this new class is either DSCP AF11 or CS1, which necessitates the previously unrequired re-marking of the Scavenger class to DSCP 0 (so that it will not be admitted into the Bulk Data class, but will fall into the Best-Effort class). Figure 5-7 illustrates the re-marking required when using this new class for Bulk Data.

Figure 5-7 Five-Class Provider-Edge Model Re-Marking Diagram



Example 5-5 shows an example CE configuration for a QoS Baseline enterprise model mapping (over a dual-T1 link) into a five-class service-provider model.

Example 5-5 CE Configuration for Five-Class Provider-Edge Model (continued)

```
!
class-map match-all ROUTING
 match ip dscp cs6
class-map match-all VOICE
 match ip dscp ef
class-map match-all INTERACTIVE-VIDEO
 match ip dscp af41
class-map match-all STREAMING-VIDEO
 match ip dscp cs4
class-map match-all MISSION-CRITICAL-DATA
 match ip dscp 25
class-map match-any CALL-SIGNALING
 match ip dscp af31
 match ip dscp cs3
class-map match-all TRANSACTIONAL-DATA
 match ip dscp af21
class-map match-all BULK-DATA
 match ip dscp af11
class-map match-all NETWORK-MANAGEMENT
 match ip dscp cs2
class-map match-all SCAVENGER
 match ip dscp cs1
!
!
policy-map CE-FIVE-CLASS-SP-MODEL
 class ROUTING
  bandwidth percent 3  ! Routing is assigned (by default) to Critical SP class
 class VOICE
  priority percent 18  ! Voice is admitted to Realtime SP class
 class INTERACTIVE-VIDEO
```

```
 priority percent 15
 set ip dscp cs5      ! Interactive-Video is assigned to the Realtime SP class
class STREAMING-VIDEO
 bandwidth percent 13
 set ip dscp af21    ! Streaming-Video is assigned to the Video SP class
class CALL-SIGNALING
 priority percent 2  ! Call-Signaling gets LLQ for this scenario
 set ip dscp cs5     ! Call-Signaling is assigned to the Realtime SP class
class MISSION-CRITICAL-DATA
 bandwidth percent 12
 random-detect
 set ip dscp af31     ! MC Data is assigned to the Critical SP class
class TRANSACTIONAL-DATA
 bandwidth percent 5
 random-detect
 set ip dscp cs3      ! Transactional Data is assigned to Critical SP class
class NETWORK-MANAGEMENT
 bandwidth percent 2  ! Net Mgmt (mainly UDP) is admitted to Video SP class
class BULK-DATA
 bandwidth percent 5  ! Bulk Data is assigned to Bulk SP class
 random-detect
class SCAVENGER
 bandwidth percent 1
 set ip dscp 0          ! Scavenger is re-marked to 0
class class-default
 bandwidth percent 24
 random-detect
!
```

Verification commands:

- **show policy**
- **show policy interface**

The **max-reserved-bandwidth** command might be required on the interface to which the preceding policy is applied.

Provider-Edge QoS Considerations

PE designs are relevant for service providers (and for enterprises that are self-managing their own MPLS VPNs). Two unique considerations for PE QoS design are discussed next:

- Service provider-to-enterprise models
- MPLS DiffServ tunneling modes

These considerations are examined in more detail in the following sections.

Service Provider-to-Enterprise Models

The PE edges facing customer CEs are complementary to the enterprise-to-service provider mapping models discussed previously. The PE designs for each class model (three, four, and five) are detailed in the following sections.

Three-Class Provider-Edge Model: PE Design

As outlined previously (and illustrated in Figure 5-15), in this model, the service provider offers three classes of service: Real-Time (strict priority, available in 5-percent increments), Critical Data (guaranteed bandwidth), and Best-Effort. The admission criterion for the Real-Time class is either DSCP EF or CS5; the admission criterion for Critical Data is DSCP CS6 (for customer routing traffic), AF31, or CS3. All other code points are re-marked to 0 by an ingress policer (not shown in this configuration example, but detailed later under the MPLS DiffServ tunneling examples). Additionally, service-provider policers can re-mark out-of-contract AF31 traffic down to AF32, which results in a higher drop preference because DSCP-based WRED is enabled on this class. As in previous examples, Example 5-6 is based on an access link of more than 3 Mbps.

Example 5-6 PE Configuration for Three-Class Provider-Edge Model

```
!
class-map match-any REALTIME
 match ip dscp ef
 match ip dscp cs5
class-map match-any CRITICAL-DATA
 match ip dscp cs6
 match ip dscp af31
 match ip dscp cs3
!
policy-map PE-THREE-CLASS-SP-MODEL
 class REALTIME
  priority percent 35          ! Realtime class gets 35% LLQ
 class CRITICAL-DATA
  bandwidth percent 40         ! Critical-Data SP class gets 40% CBWFQ
  random-detect dscp-based     ! DSCP-based WRED enabled on class
 class class-default
  fair-queue                   ! Best Effort SP class gets FQ
  random-detect                ! WRED enabled on Best Effort SP class
!
```

Verification commands:

- **show policy**
- **show policy interface**

Four-Class Provider-Edge Model: PE Design

Building on the previous model (and as illustrated in Figure 5-6), a fourth class is added to this SP model, which can be used for either Bulk Data or Streaming-Video. The admission criterion for this new class is either DSCP AF21 or CS2. Out-of-contract AF21 traffic offered to this class can be marked down to AF22. In this particular example, the class is being called Video, but it is important to keep in mind that the customer can offer any traffic desired to this class, provided that it is marked appropriately. For this reason (although it normally is not required on UDP-based flows such as Streaming-Video), DSCP-based WRED is enabled on this class to aggressively drop out-of-contract traffic as needed. As in previous examples, Example 5-7 is based on an access link of more than 3 Mbps.

Example 5-7 PE Configuration for Four-Class Provider-Edge Model

```
!
class-map match-any REALTIME
 match ip dscp ef
 match ip dscp cs5
class-map match-any CRITICAL-DATA
```

```
 match ip dscp cs6
 match ip dscp af31
 match ip dscp cs3
class-map match-any VIDEO
 match ip dscp af21
 match ip dscp cs2
!
policy-map PE-FOUR-CLASS-SP-MODEL
 class REALTIME
  priority percent 35          ! Realtime SP class gets 35% LLQ
 class CRITICAL-DATA
  bandwidth percent 25         ! Critical-Data SP class gets 40% CBWFQ
  random-detect dscp-based     ! DSCP-based WRED enabled on class
 class VIDEO
  bandwidth percent 15         ! Video SP class gets 15% CBWFQ
  random-detect dscp-based     ! DSCP-based WRED enabled on "Video" SP class
 class class-default
  fair-queue                   ! Best Effort SP class gets FQ
  random-detect                ! WRED enabled on Best Effort SP class
!
```

Verification commands:

- **show policy**
- **show policy interface**

Five-Class Provider-Edge Model: PE Design

Building again on the previous model (and as illustrated in Figure 5-7), a fifth class is added that can be used for either Bulk Data or Video (whichever wasn't used under the four-class model). In this example, the new class is used for Bulk Data. The admission criterion for this new class is either DSCP AF11 or CS1. Out-of-contract AF11 traffic offered to this class can be re-marked to AF12 and can be discarded earlier by the DSCP-based WRED algorithm operating on the output queue for this class.

To prevent long TCP sessions of the Bulk Data SP class from dominating bandwidth intended for the Best-Effort class, a bandwidth guarantee is offered to the Best-Effort class. This guarantee might require the use of the **max-reserved-bandwidth** override under the applied interface configuration. As in the previous examples, an access link of more than 3 Mbps is assumed in Example 5-8.

Example 5-8 PE Configuration for Five-Class Provider-Edge Model

```
!
class-map match-any REALTIME
 match ip dscp ef
 match ip dscp cs5
class-map match-any CRITICAL-DATA
 match ip dscp cs6
 match ip dscp af31

 match ip dscp cs3
class-map match-any VIDEO
 match ip dscp af21
 match ip dscp cs2
class-map match-any BULK-DATA
 match ip dscp af11
 match ip dscp cs1
!
policy-map PE-FIVE-CLASS-SP-MODEL
 class REALTIME
  priority percent 35          ! Realtime SP class gets 35% LLQ
```

```
 class CRITICAL-DATA
  bandwidth percent 20         ! Critical-Data SP class gets 40% CBWFQ
  random-detect dscp-based     ! DSCP-based WRED enabled on class
 class VIDEO
  bandwidth percent 15         ! Video SP class gets 15% CBWFQ
  random-detect dscp-based     ! DSCP-based WRED enabled on "Video" SP class
 class BULK-DATA
  bandwidth percent 5          ! Bulk Data SP class gets 15% CBWFQ
  random-detect dscp-based     ! DSCP-based WRED enabled on Bulk Data SP class
 class class-default
  bandwidth percent 25         ! Best Effort SP class gets 25% CBWFQ
  random-detect                ! WRED enabled on Best Effort SP class
!
```

Verification commands:

- **show policy**
- **show policy interface**

MPLS DiffServ Tunneling Modes

As described in previous examples, some service providers re-mark packets at Layer 3 to indicate whether traffic is in contract or out-of-contract. Although this conforms to DiffServ standards, such as RFC 2597, this is not always desirable from an enterprise customer's standpoint.

Because MPLS labels include 3 bits that commonly are used for QoS marking, it is possible to "tunnel DiffServ"—that is, preserve Layer 3 DiffServ markings through a service provider's MPLS VPN cloud while still performing re-marking (via MPLS EXP bits) within the cloud to indicate in- or out-of-contract traffic.

RFC 3270 defines three distinct modes of MPLS DiffServ tunneling; each is discussed in detail in the following sections:

- Uniform Mode
- Short Pipe Mode
- Pipe Mode

Uniform Mode

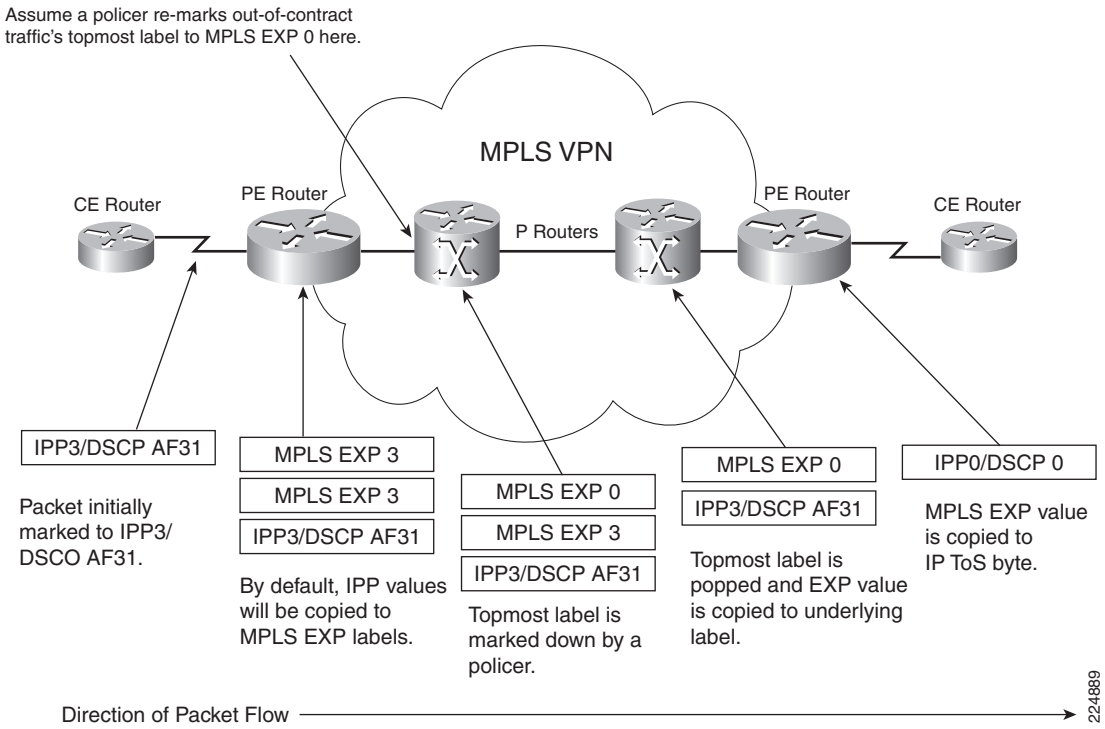
Uniform Mode generally is utilized when the customer and service provider share the same DiffServ domain, as in the case of an enterprise deploying its own MPLS VPN core.

In Uniform Mode, which is the default mode, the first 3 bits of the IP ToS field (IP Precedence bits) automatically are mapped to the MPLS EXP bits on the ingress PE as labels are pushed onto the packets.

If policers or any other mechanisms re-mark the MPLS EXP values within the MPLS core, these marking changes are propagated to lower-level labels and eventually are propagated to the IP ToS field (MPLS EXP bits are mapped to IP Precedence values on the egress PE).

Figure 5-8 shows the behavior of Uniform Mode MPLS DiffServ tunneling.

Figure 5-8 MPLS DiffServ Uniform Tunneling Mode Operation

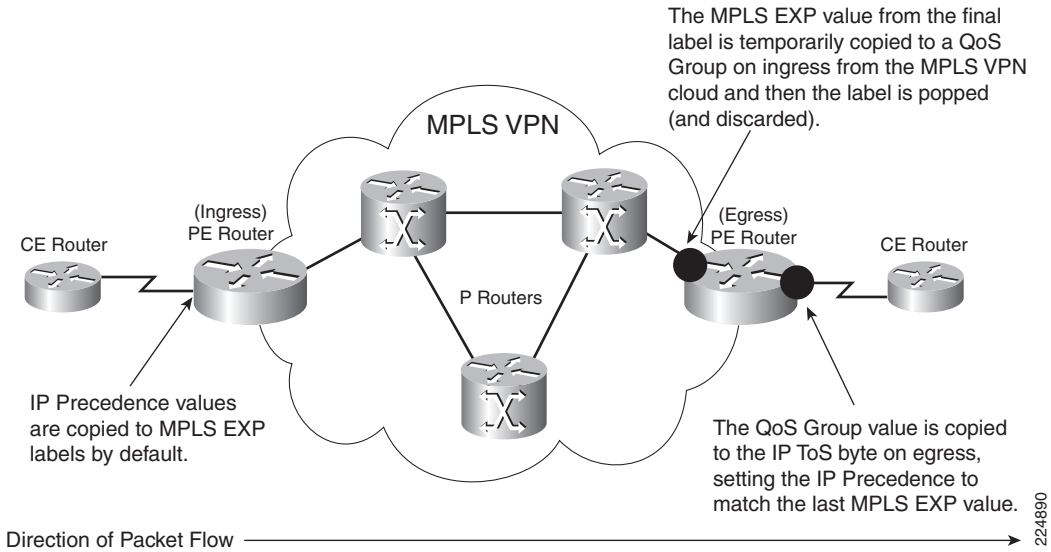


The mapping of IP Precedence to MPLS EXP is performed by default on PEs for customer-to-provider traffic.

However, for provider-to-customer egress traffic (from the MPLS VPN cloud), additional configuration is required on the PE to achieve mapping of MPLS EXP to IP Precedence. This is because the final label is popped (and discarded) when it is received from the MPLS VPN cloud and, therefore, cannot be used as a match criterion for policies applied to the egress interface of the final PE router (facing the destination CE). The solution is to copy the final MPLS EXP bit values to a temporary placeholder on PE ingress from the MPLS core (before the label is discarded) and then use these temporary placeholder values for setting the IP Precedence bits on egress to the customer CE.

Cisco IOS provides two such temporary placeholders, the QoS Group and the Discard Class. For Uniform Mode scenarios, it is recommended to copy the MPLS EXP values to QoS Group values on ingress from the MPLS VPN cloud. (The Discard Class is recommended for use in Pipe Mode scenarios only.) Then QoS Group values can be copied to IP Precedence values (on egress to the customer CE). Figure 5-9 illustrates the policies required for a single direction for Uniform Mode MPLS DiffServ tunneling. (This policy also would be required on the complementary interfaces for the reverse traffic direction.)

Figure 5-9 MPLS DiffServ Uniform Tunneling Mode Policies



Example 5-9 shows the configuration for Uniform Mode operation on a PE.

Example 5-9 PE Configuration for MPLS DiffServ Uniform Mode Tunneling

```
!
policy-map MPLSEXP-TO-QOSGROUP
  class class-default
   set qos-group mpls experimental topmost  ! Copies MPLS EXP to QoS Group
!
 policy-map QOSGROUP-TO-IPP
  class class-default
   set precedence qos-group                 ! Copies QoS Group to IPP
!
…
!
interface ATM2/0
 no ip address
 no atm ilmi-keepalive
!
interface ATM2/0.1 point-to-point
 description ATM-OC3 TO MPLS VPN CORE       ! Link to/from MPLS VPN Core
 ip address 20.2.34.4 255.255.255.0
 pvc 0/304
  vbr-nrt 149760 149760
  service-policy input MPLSEXP-TO-QOSGROUP  ! MPLS EXP to QoS Group on ingress
 !
 tag-switching ip
!
…
!
interface FastEthernet1/0
 description FE TO CUSTOMER RED CE         ! Link to/from CE
 ip vrf forwarding RED
 ip address 10.1.45.4 255.255.255.0
 service-policy output QOSGROUP-TO-IPP     ! QoS Group to IPP on egress to CE
!
```

Verification commands:

- **show policy**
- **show policy interface**

Of course, additional QoS policies (to these Uniform Mode tunneling policies), such as queuing or WRED, can be applied on the PE-to-CE egress link (as detailed earlier in the previous section).

Short Pipe Mode

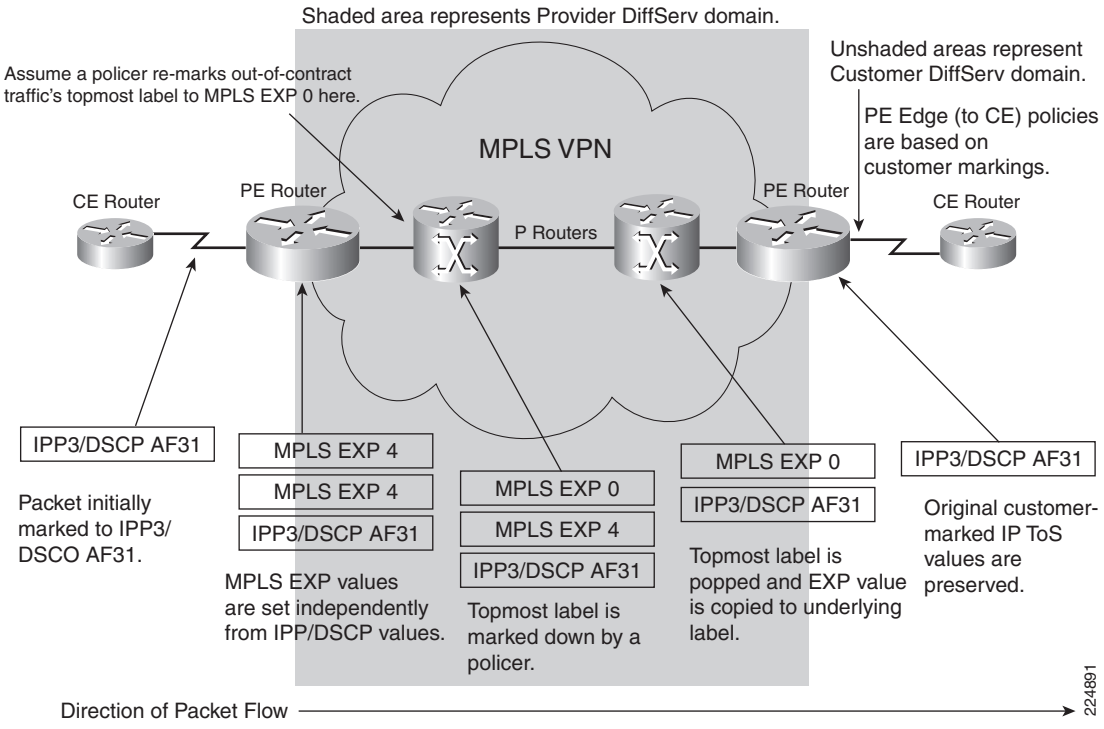
Short Pipe Mode is utilized when the customer and service provider are in different DiffServ domains. (The service provider's DiffServ domain begins at the ingress PE's ingress interface and terminates on the egress PE's ingress interface.)

This mode is useful when the service provider wants to enforce its own DiffServ policy and the customer requests that its DiffServ information be preserved through the MPLS VPN cloud. Short Pipe Tunneling Mode provides DiffServ transparency through the service provider network (as does Pipe Mode).

The outmost label is utilized as the single most meaningful information source as it relates to the service provider's QoS PHB. On MPLS label imposition, the IP classification is not copied into the outermost label's EXP. Instead, the value for the MPLS EXP is set explicitly on the ingress PE's ingress interface, according to the service provider's administrative policies.

In the case of any re-marking occurrence within the service provider's MPLS VPN cloud, changes are limited to MPLS EXP re-marking only and are not propagated down to the underlying IP packet's ToS byte. Figure 5-10 shows the operation of Short Pipe Mode MPLS DiffServ tunneling.

Figure 5-10 MPLS DiffServ Short Pipe Mode Tunneling Operation



MPLS EXP values can be marked in any way that the provider wants to provide local significance. Figure 5-11 shows an example use of MPLS EXP markings to indicate in- or out-of-contract traffic for a five-class service-provider model.

Figure 5-11 Five-Class Service Provider Model Short Pipe Mode Re-Marking Diagram

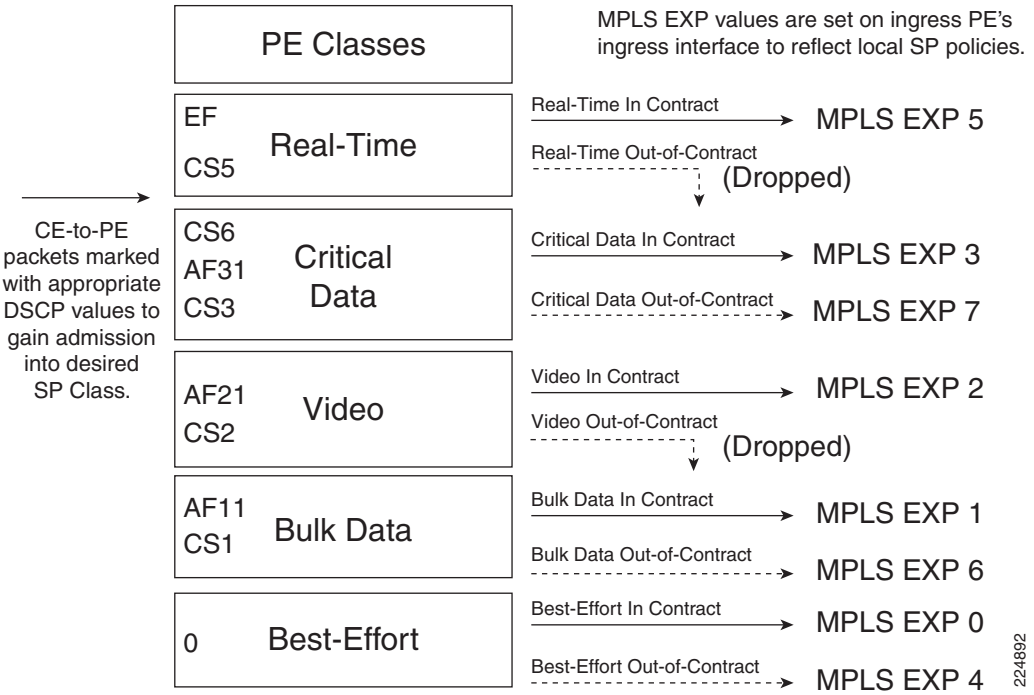
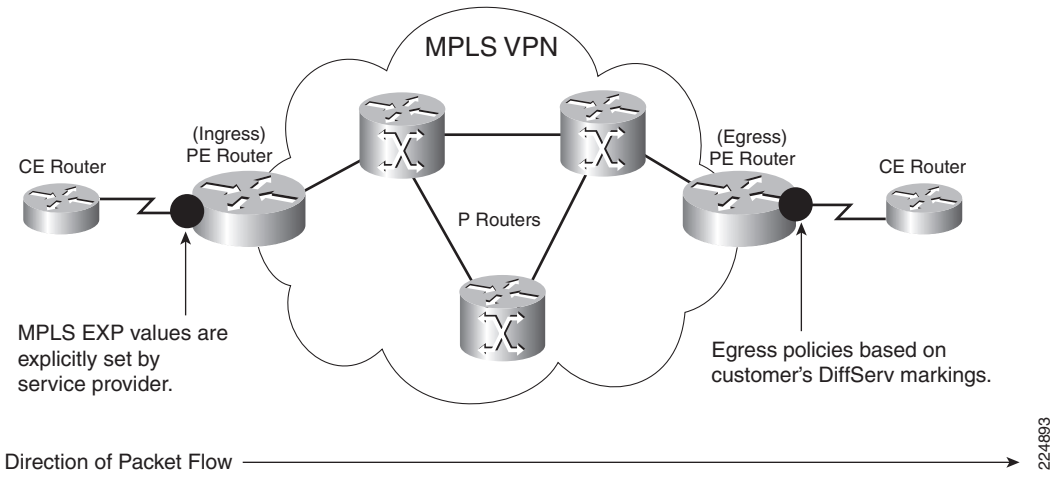


Figure 5-12 shows the ingress PE ingress interface re-marking policies for Short Pipe Mode, based on the re-marking diagram provided in Figure 5-11. No mapping from MPLS EXP to QoS Group is needed on the egress PE's ingress interface (as was required for Uniform Mode) because the MPLS EXP value loses relevance beyond this interface.

Any egress policies on the egress PE's egress interface (facing the customer's destination CE), are based on IP Precedence or DSCP values (which have remained untouched). This is the main difference between Short Pipe Mode and Pipe Mode.

Figure 5-12 shows the interfaces in which explicit policy configuration is required for Short Pipe Mode MPLS DiffServ tunneling.

Figure 5-12 MPLS DiffServ Short Pipe Mode Tunneling Policies



Example 5-10 shows the configuration for Short Pipe Mode operation on a PE. Traffic received from CEs is marked explicitly (through MPLS EXP values) to reflect the service provider's policies. In this example, the customer is given a 3-Mbps CIR through an FE access link. The provider is using a five-class model with 35 percent for Real-Time traffic, 20 percent for Critical Data traffic, 15 percent for Video traffic, 5 percent for Bulk Data traffic, and 25 percent for Best-Effort traffic. On PE-to-CE links (in the egress direction), queuing and dropping policies based on customer IP DiffServ markings also are recommended (as was discussed previously).

Example 5-10 PE Configuration for MPLS DiffServ Short Pipe Mode Tunneling (continued)

```
!
class-map match-any REALTIME
  match ip dscp ef
  match ip dscp cs5
class-map match-any CRITICAL-DATA
  match ip dscp cs6
  match ip dscp af31
  match ip dscp cs3
class-map match-any VIDEO
  match ip dscp af21
  match ip dscp cs2
class-map match-any BULK-DATA
  match ip dscp af11
  match ip dscp cs1
!
!
 policy-map PE-FIVE-CLASS-SHORT-PIPE-MARKING
  class REALTIME
   police cir 1050000
     conform-action set-mpls-exp-topmost-transmit 5  ! Conforming RT set to 5
     exceed-action drop                              ! Excess Realtime is dropped
  class CRITICAL-DATA
   police cir 600000
     conform-action set-mpls-exp-topmost-transmit 3  ! Critical Data set to 3
     exceed-action set-mpls-exp-topmost-transmit 7   ! Excess Critical set 7
  class VIDEO
   police cir 450000
     conform-action set-mpls-exp-topmost-transmit 2  ! Conforming Video set to 2
     exceed-action drop                              ! Excess Video dropped
  class BULK-DATA
   police cir 150000
     conform-action set-mpls-exp-topmost-transmit 1  ! Conforming Bulk set to 1
     exceed-action set-mpls-exp-topmost-transmit 6   ! Excess Bulk set to 6
  class class-default
   police cir 750000
     conform-action set-mpls-exp-topmost-transmit 0  ! Conforming BE set to 0
     exceed-action set-mpls-exp-topmost-transmit 4   ! Excess BE set to 4
!
…
!
interface FastEthernet1/0
 description FE TO CUSTOMER RED CE                   ! Link to/from CE
 ip vrf forwarding RED
 ip address 10.1.12.2 255.255.255.0
 service-policy input PE-FIVE-CLASS-SHORT-PIPE-MARKING
!
```

Verification commands:

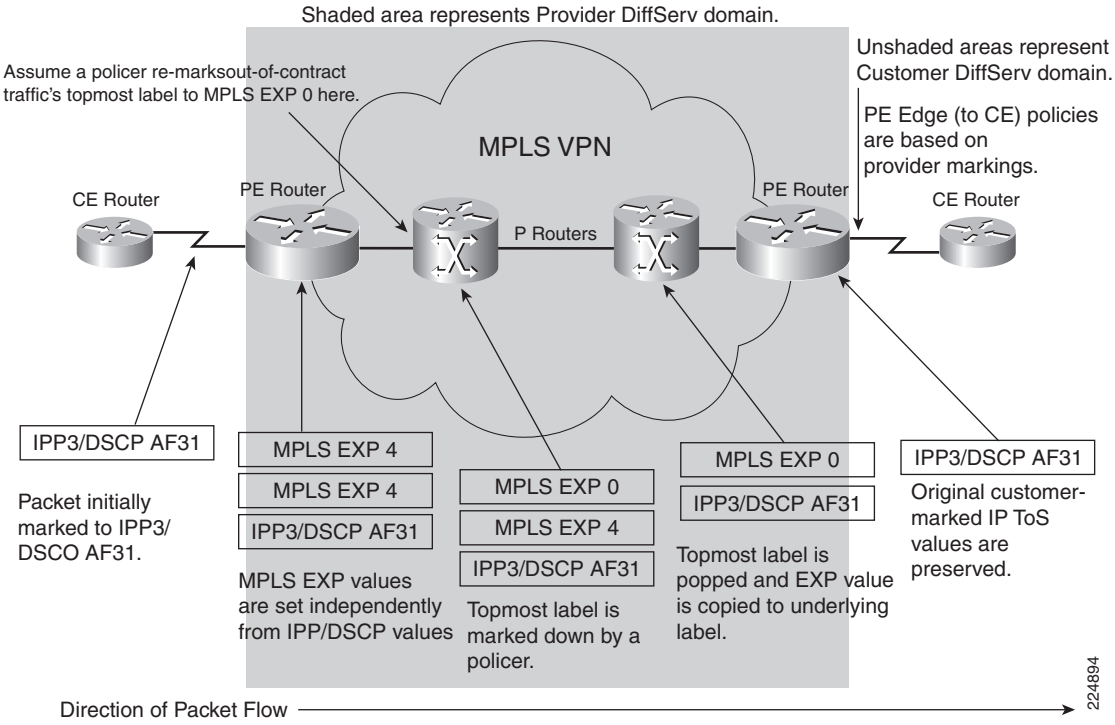
- **show policy**
- **show policy interface**

Pipe Mode

The main difference between Short Pipe Mode and Pipe Mode MPLS DiffServ tunneling is that the PE egress policies (toward the customer CEs) are provisioned according to the *service provider's* explicit markings and re-markings, not the enterprise customer's IP DiffServ markings (although these are preserved). As with Short Pipe Mode, any changes to label markings that occur within the service provider's cloud do not get propagated to the IP ToS byte when the packet leaves the MPLS network.

Because egress PE-to-CE QoS policies in Pipe Mode are dependent on the last MPLS EXP value, this value must be preserved before the final label is popped. A temporary placeholder (as used in Uniform Mode operation) is again required. On the final PE router in a given path, the MPLS EXP value is copied to the QoS Group value. Optionally, a Discard Class value also might set drop preference at the same time. Thereafter, egress queuing or dropping policies are performed based on these QoS Group/Discard Class values. Figure 5-13 illustrates the Pipe Mode MPLS DiffServ tunneling operation.

Figure 5-13 MPLS DiffServ Pipe Mode Tunneling Operation



QoS Groups and Discard Classes can be combined to provide virtual DiffServ PHB classification. For example, RFC 2597 assured-forwarding PHBs can be mimicked using QoS Group values 1 through 4 (to represent the AF class) coupled with Discard Class values 1 through 3 (to represent the drop preference). In general, QoS Group and Discard Class values are arbitrary and have only local significance. However, an exception is found when WRED is configured to selectively drop based on Discard Class values, in which case the lower Discard Class values are dropped first (by default). If no Discard Class value is assigned explicitly, the value defaults to 0.

Figure 5-14 shows the points where policies are required for Pipe Mode MPLS DiffServ tunneling.

Figure 5-14 MPLS DiffServ Pipe Mode Tunneling Policies

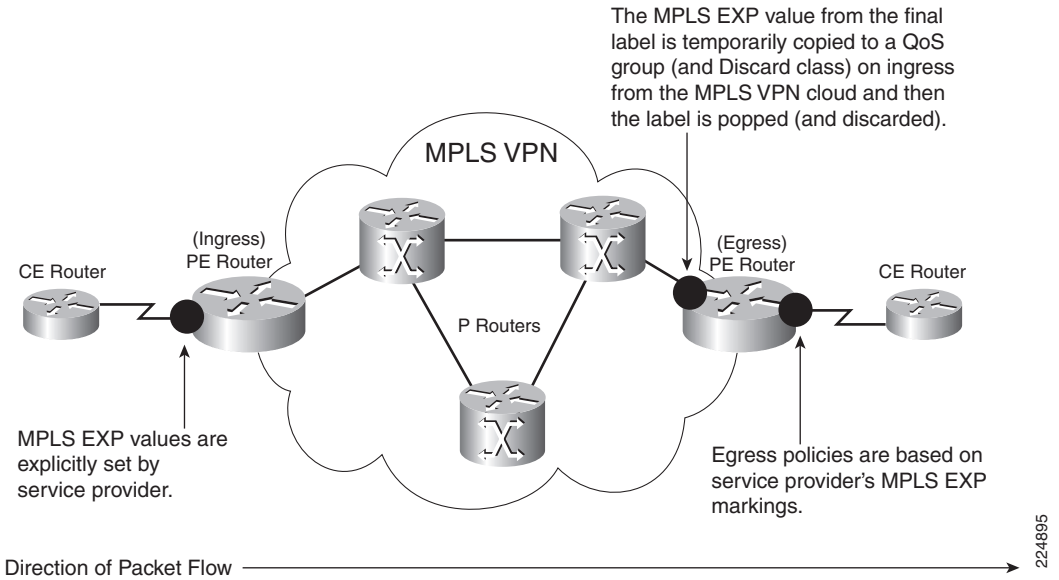
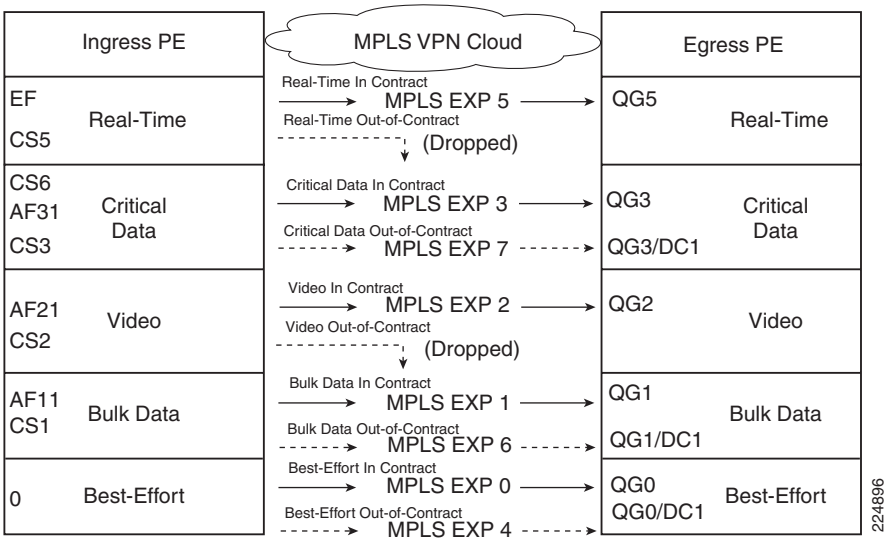


Figure 5-15 illustrates adapting the five-class service provider model to Pipe Mode. The first set of re-markings shows ingress PE re-marking from DSCP to MPLS EXP values, depending on whether the traffic is in contract or out-of-contract. The second set of markings shows how these MPLS EXP values can be mapped to QoS Groups (QG) and Discard Classes (DC) to provide PHB classification and provisioning on PE-to-CE links (without altering the IP DSCP values of the tunneled packets).

Example 5-11 shows the configuration for bidirectional re-marking on a PE router to support Pipe Mode operation. Traffic received from CEs is marked explicitly (through MPLS EXP values) to reflect the service provider's policies. Then traffic (traversing in the opposite direction) received from the MPLS VPN core is mapped to QoS Groups and Discard Classes so that PE-to-CE PHB egress policies can be performed against provider re-markings. In this example, the customer has contracted for 3-Mbps service over an FE link. Hierarchical policies are used to achieve queuing within (3 Mbps) shaping over this (100-Mbps) link. Additionally, Discard-class WRED is enabled on the output queues so that dropping decisions are based on Discard-class values (not IP ToS or DSCP values). Furthermore, Discard-class dropping thresholds are tuned so that Discard-Class 1 (indicating out-of-contract traffic) is dropped more aggressively than Discard-Class 0 (mimicking DSCP-based WRED behavior), which is more consistent with RFC 2597 Assured-Forwarding PHBs.

Figure 5-15 Five-Class Service Provider Model Pipe Mode Ingress and Egress Re-Marking Diagram



Example 5-11 PE Configuration for MPLS DiffServ Pipe Mode Tunneling

```
!
class-map match-any REALTIME
  match ip dscp ef
  match ip dscp cs5
class-map match-any CRITICAL-DATA
  match ip dscp cs6
  match ip dscp af31
  match ip dscp cs3
class-map match-any VIDEO
  match ip dscp af21
  match ip dscp cs2
class-map match-any BULK-DATA
  match ip dscp af11
  match ip dscp cs1
!
!
 class-map match-all MPLS-EXP-7
  match mpls experimental topmost 7          ! Matches MPLS EXP 7
 class-map match-all MPLS-EXP-6
  match mpls experimental topmost 6          ! Matches MPLS EXP 6
 class-map match-all MPLS-EXP-5
  match mpls experimental topmost 5          ! Matches MPLS EXP 5
 class-map match-all MPLS-EXP-4
  match mpls experimental topmost 4          ! Matches MPLS EXP 4
 class-map match-all MPLS-EXP-3
  match mpls experimental topmost 3          ! Matches MPLS EXP 3
 class-map match-all MPLS-EXP-2
  match mpls experimental topmost 2          ! Matches MPLS EXP 2
 class-map match-all MPLS-EXP-1
  match mpls experimental topmost 1          ! Matches MPLS EXP 1
 class-map match-all MPLS-EXP-0
  match mpls experimental topmost 0          ! Matches MPLS EXP 0
!
!
 class-map match-all QOSGROUP5
  match qos-group 5                          ! Matches QoS Group 5
```

```
 class-map match-all QOSGROUP3
  match qos-group 3                              ! Matches QoS Group 3
 class-map match-all QOSGROUP2
  match qos-group 2                              ! Matches QoS Group 2
 class-map match-all QOSGROUP1
  match qos-group 1                              ! Matches QoS Group 1
 class-map match-all QOSGROUP0
  match qos-group 0                              ! Matches QoS Group 0
!
!
 policy-map PIPE-MARKING                          ! Sets MPLS EXP Values
  class REALTIME
   police cir 1050000
     conform-action set-mpls-exp-topmost-transmit 5  ! Conforming RT set to 5
     exceed-action drop                              ! Excess Realtime is dropped
  class CRITICAL-DATA
   police cir 600000
     conform-action set-mpls-exp-topmost-transmit 3 ! Critical Data set to 3
     exceed-action set-mpls-exp-topmost-transmit 7  ! Excess Critical set 7
  class VIDEO
   police cir 450000
     conform-action set-mpls-exp-topmost-transmit 2  ! Conforming Video set to 2
     exceed-action drop                              ! Excess Video dropped
  class BULK-DATA
   police cir 150000
     conform-action set-mpls-exp-topmost-transmit 1  ! Conforming Bulk set to 1
     exceed-action set-mpls-exp-topmost-transmit 6  ! Excess Bulk set to 6
  class class-default
   police cir 750000
     conform-action set-mpls-exp-topmost-transmit 0  ! Conforming BE set to 0
     exceed-action set-mpls-exp-topmost-transmit 4   ! Excess BE set to 4
!
!
 policy-map MPLSEXP-QOSGROUP-DISCARDCLASS   ! Maps MPLS EXP to QG/DC values
  class MPLS-EXP-5
   set qos-group 5            ! Conforming Realtime is set to QG 5
  class MPLS-EXP-3
   set qos-group 3            ! Conforming Critical Data is set to QG 3

  class MPLS-EXP-7
   set qos-group 3            ! Excess Critical Data is set to QG3
   set discard-class 1        ! Excess Critical Data has DC set to 1
  class MPLS-EXP-2
   set qos-group 2            ! Conforming Video is set to QG 2
  class MPLS-EXP-1
   set qos-group 1            ! Conforming Bulk is set to QG 1
  class MPLS-EXP-6
   set qos-group 1            ! Excess Bulk is set to QG 1
   set discard-class 1        ! Excess Bulk has DC set to 1
  class MPLS-EXP-0
   set qos-group 0            ! Conforming Best Effort is set to QG 0
  class MPLS-EXP-4
   set qos-group 0            ! Excess Best Effort is set to QG 0
   set discard-class 1        ! Excess Best Effort has DC set to 1
!
!
 policy-map PE-CE-QUEUING     ! Queuing policy for PE to CE link
  class QOSGROUP5
   priority percent 35        ! Voice class gets 35% LLQ
  class QOSGROUP3
   bandwidth percent 20        ! Critical Data class gets 20% CBWFQ
   random-detect discard-class-based                ! DC-Based WRED is enabled
   random-detect discard-class 0   30    40    10    ! DC 0 is tuned for WRED
   random-detect discard-class 1   20    40    10    ! DC 1 is tuned for WRED
```

```
  class QOSGROUP2
   bandwidth percent 15       ! Video class gets 15% CBWFQ
  class QOSGROUP1
   bandwidth percent 5        ! Bulk class gets 5% CBWFQ
   random-detect discard-class-based                ! DC-Based WRED is enabled
   random-detect discard-class 0   30    40    10   ! DC 0 is tuned for WRED
   random-detect discard-class 1   20    40    10   ! DC 1 is tuned for WRED
  class QOSGROUP0
   bandwidth percent 25    ! Best Effort class gets 25% CBWFQ
   random-detect discard-class-based                ! DC-Based WRED is enabled
   random-detect discard-class 0   30    40    10   ! DC 0 is tuned for WRED
   random-detect discard-class 1   20    40    10   ! DC 1 is tuned for WRED
!
!
policy-map PE-CE-SHAPING-QUEUING     ! Customer has 3 Mbps CIR over FE
  class class-default
   shape average 3000000             ! Shaping policy for 3 Mbps CIR
   service-policy PE-CE-QUEUING      ! Nested queuing policy
!
interface ATM2/0
 no ip address
 no atm ilmi-keepalive
!
interface ATM2/0.1 point-to-point
 description ATM-OC3 TO MPLS VPN CORE       ! Link to/from MPLS VPN Core
 ip address 20.2.34.4 255.255.255.0
 pvc 0/304
  vbr-nrt 149760 149760
  service-policy input MPLSEXP-QOSGROUP-DISCARDCLASS ! MPLS EXP to QG/DC
 !
 tag-switching ip
!
!
interface FastEthernet1/0
 description FE TO CUSTOMER RED CE                  ! Link to/from CE
 ip vrf forwarding RED
 ip address 10.1.12.2 255.255.255.0
 service-policy input PIPE-MARKING                  ! Pipe marking policy
 service-policy output PE-CE-SHAPING-QUEUING        ! Shaping/Queuing policy
!
```

Verification commands:

- **show policy**
- **show policy interface**

Pipe Mode with an Explicit Null LSP

When CEs are provider managed, some providers prefer to offload the ingress MPLS EXP marking of customer traffic from the PE and push these policies out to the ingress interface of the CE. However, because the CE-to-PE link is regular IP (not MPLS), a difficulty arises as to how to set the provider's marking without affecting the IP DiffServ markings that the customer has set (because, again, these are to be preserved and untouched through the MPLS VPN cloud in Pipe Mode operation). Therefore, a solution to this scenario was introduced in Cisco IOS Release 12.2(13)T, with the Pipe Mode MPLS DiffServ tunneling with an Explicit Null LSP feature.

This feature prepends an Explicit Null LSP label for customer traffic headed from the CE to the PE. This label is not used for MPLS switching; it is used only to preserve the provider's MPLS EXP markings over the CE-to-PE link. On the PE, the MPLS EXP values are copied to regular MPLS labels that are pushed onto the packet (which are used for MPLS switching), and the explicit null label is discarded.

Thus, the ingress marking policies from the PE are pushed to the managed CE. This expands the provider's DiffServ domain to include the (managed) CEs. All other aspects of Pipe Mode operation and configuration, however, remain the same.

Figure 5-16 MPLS DiffServ Pipe Mode with an Explicit Null LSP Tunneling Operation

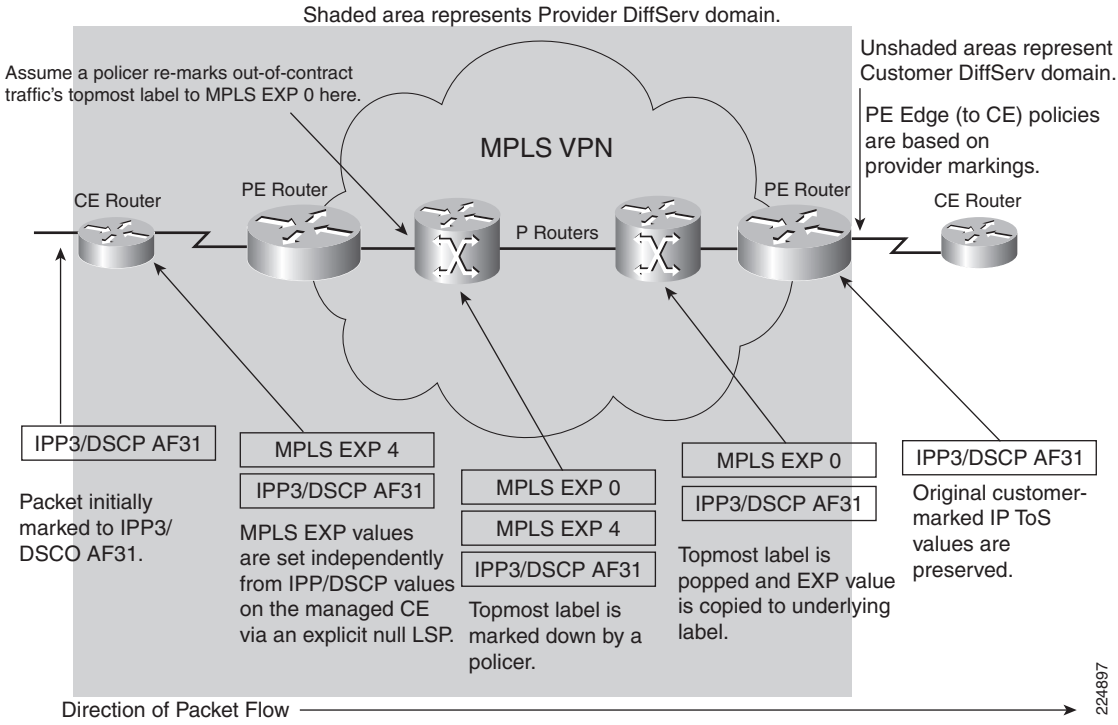
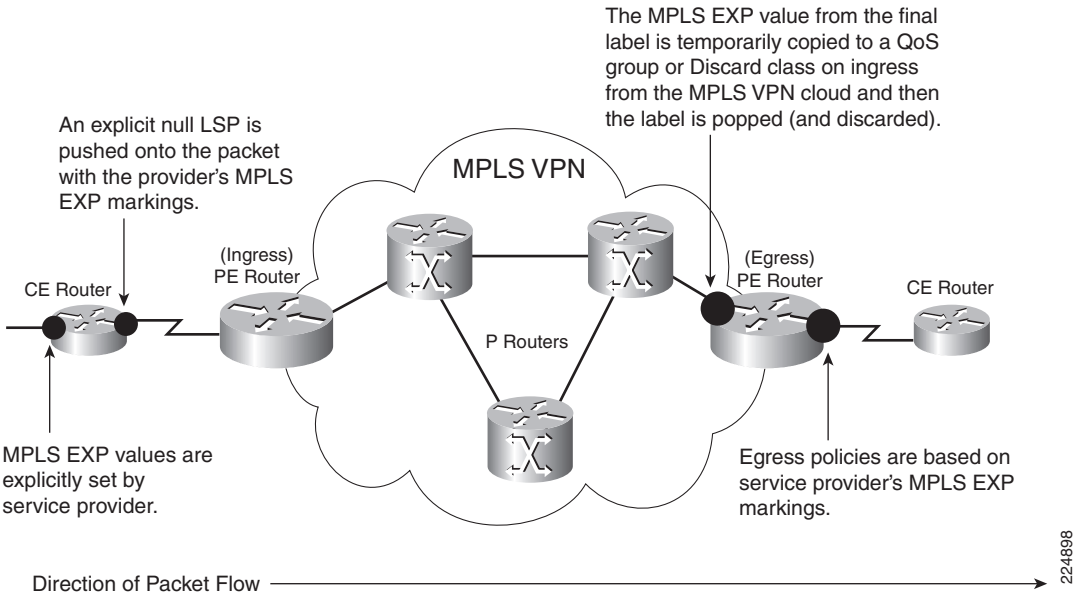


Figure 5-17 shows the points where policies are required for Pipe Mode with Explicit Null LSP MPLS DiffServ tunneling.

As noted, PE configurations remain the same as with normal Pipe Mode, with the exception that ingress MPLS EXP marking policies have been removed. These policies now are set on the managed CE, as shown in Example 5-12. The provider has contracted for a CIR of 3 Mbps in a five-class model. All ingress traffic is policed on the customer edge of the CE and is marked (through MPLS EXP values) to indicate whether it is in contract or out-of-contract. Then an Explicit Null LSP is pushed onto the packet to carry these MPLS EXP markings from the CE to the PE. Optionally, queuing policies can be added on the CE-to-PE link, but for simplicity, these have been omitted from this example because they already have been covered.

Figure 5-17 MPLS DiffServ Pipe Mode with an Explicit Null LSP Tunneling Policies



Example 5-12 Managed CE Configuration for MPLS DiffServ Pipe Mode with an Explicit Null LSP Tunneling

```
!
class-map match-any REALTIME
  match ip dscp ef
  match ip dscp cs5
class-map match-any CRITICAL-DATA
  match ip dscp cs6
  match ip dscp af31
  match ip dscp cs3
class-map match-any VIDEO
  match ip dscp af21
  match ip dscp cs2
class-map match-any BULK-DATA
  match ip dscp af11
  match ip dscp cs1
!
!
 policy-map PIPE-EXPLICIT-NULL-MARKING
  class REALTIME
   police cir 1050000
     conform-action set-mpls-exp-topmost-transmit 5  ! Conforming RT set to 5
     exceed-action drop                              ! Excess Realtime is dropped

  class CRITICAL-DATA
   police cir 600000
     conform-action set-mpls-exp-topmost-transmit 3  ! Critical Data set to 3
     exceed-action set-mpls-exp-topmost-transmit 7   ! Excess Critical set 7
  class VIDEO
   police cir 450000
     conform-action set-mpls-exp-topmost-transmit 2  ! Conforming Video set to 2
     exceed-action drop                              ! Excess Video dropped
  class BULK-DATA
   police cir 150000
     conform-action set-mpls-exp-topmost-transmit 1  ! Conforming Bulk set to 1
     exceed-action set-mpls-exp-topmost-transmit 6   ! Excess Bulk set to 6
```

```
  class class-default
   police cir 750000
     conform-action set-mpls-exp-topmost-transmit 0  ! Conforming BE set to 0
     exceed-action set-mpls-exp-topmost-transmit 4   ! Excess BE set to 4
!
…
!
interface FastEthernet0/0
 description FE to Customer Network                 ! Link to/from customer
 ip address 10.1.1.1 255.255.255.0
 service-policy input PIPE-EXPLICIT-NULL-MARKING    ! MPLS EXP set on ingress
!
!
interface FastEthernet0/1
 description FE TO PE                                ! Link to/from PE
 ip address 10.1.12.1 255.255.255.0
 duplex auto
 speed auto
 mpls ip encapsulate explicit-null                   ! Explicit Null LSP is added
!
```

Verification commands:

- **show policy**
- **show policy interface**

Note For a comprehensive case study example of MPLS VPN QoS designs, refer to Figure 15-22 and Examples 15-29 through 15-32 of the Cisco Press book, *End-to-End QoS Network Design* by Tim Szigeti and Christina Hattingh.

Summary

MPLS VPNs are rapidly gaining popularity as private WAN alternatives. This chapter presented QoS design principles and designs to achieve end-to-end service levels over MPLS VPNs. The foremost design principle is that enterprise subscribers and service providers have to cooperatively deploy QoS over MPLS VPNs in a consistent and complementary manner.

Enterprise (customer) considerations, such as class-collapsing guidelines and traffic-mixing principles, were overviewed along with re-marking examples.

Service-provider edge QoS policies were presented for three-, four-, and five-class edge models. Additionally, details on how RFC 3270 tunneling modes (Uniform, Short Pipe, and Pipe) can be implemented within Cisco IOS Software were provided.

References

Standards

- RFC 2597 "Assured Forwarding PHB Group"
  http://www.ietf.org/rfc/rfc2597
- RFC 2547, “BGP/MPLS VPNs”

http://www.ietf.org/rfc/rfc2547.txt

- RFC 2702, "Requirements for Traffic Engineering over MPLS"

  http://www.ietf.org/rfc/rfc2702.txt
- RFC 2917, "A Core MPLS IP VPN Architecture"

  http://www.ietf.org/rfc/rfc2917.txt
- RFC 3270, "Multiprotocol Label Switching (MPLS) Support of Differentiated Services"

  http://www.ietf.org/rfc/rfc3270.txt

Books

- Alwayn, Vivek. *Advanced MPLS Design and Implementation.* Indianapolis: Cisco Press, 2001.
- Pepelnjak, Ivan, and Jim Guichard. *MPLS and VPN Architectures.* Cisco Press, 2002.
- Pepelnjak, Ivan, Jim Guichard, and Jeff Apcar. *MPLS and VPN Architectures*. Cisco Press, 2003.
- Osborne, Eric, and Ajay Simha. *Traffic Engineering with MPLS.* Cisco Press, 2003.
- Szigeti, Tim and Christina Hattingh. *End-to-End QoS Network Design: Quality of Service in LANs, WANs and VPNs*. Indianapolis: Cisco Press, 2004.

Cisco Documentation

- Configuring MPLS and MPLS traffic engineering (Cisco IOS Release 12.2)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fswtch_c/swprt3/xcftagc.htm
- MPLS VPNS (Cisco IOS Release 12.2.13T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftvpn13.htm
- MPLS DiffServ tunneling modes (Cisco IOS Release 12.2.13T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftdtmode.htm
- MPLS DiffServ-Aware Traffic Engineering (DS-TE) (Cisco IOS Release 12.2.4T)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t4/ft_ds_te.htm
- MPLS Cisco IOS documentation main link (Cisco IOS Release 12.3)

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123cgcr/swit_vcg.htm#999526