QoS

Implementing Cisco Quality of Service

Volume 3

Version 2.3

Student Guide

Text Part Number: 97-2812-01

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Table of Contents

Volume 3

Module 9

QoS Best Practices

Overview

IP was designed to provide best-effort service for delivery of data packets and to run across virtually any network transmission media and system platform. To manage applications such as voice over IP, streaming video, e-commerce, enterprise resource planning (ERP) applications, and others, a network requires quality of service (QoS). Different applications have varying needs for delay, delay variation (jitter), bandwidth, packet loss, and availability. These parameters form the basis of QoS. The IP network should be designed to provide the requisite QoS to applications.

To facilitate true end-to-end QoS on an IP network, the Internet Engineering Task Force (IETF) has defined two models: Integrated Services (IntServ) and Differentiated Services (DiffServ). IntServ follows the signaled QoS model, in which the end-hosts signal their QoS need to the network. DiffServ works on the provisioned QoS model, in which network elements are set up to service multiple classes of traffic with varying QoS requirements.

This module describes the implementation of the DiffServ model in service provider and enterprise networks. The first lesson describes baseline traffic classifications. The second lesson describes a DiffServ implementation in a typical enterprise campus and service provider network. The third lesson describes how QoS tools can mitigate DoS attacks through the use of control plane, data plane, and Network-Based Application Recognition (NBAR) known-worm policing.

Module Objectives

Upon completing this module, you will be able to correctly select the most appropriate QoS mechanisms for providing QoS using Cisco best practices in service provider and enterprise networks. This ability includes being able to meet these objectives:

- Describe the set of classification practices that most closely represent Cisco QoS best practices
- Describe the set of QoS mechanisms used to implement Cisco end-to-end QoS best practices in a typical enterprise network connected through a service provider that is providing Layer 3 IP services
- Describe steps recommended to mitigate DoS attacks and worm attacks using QoS tools

Lesson 1

Understanding Traffic Classification Best Practices

Overview

Traffic classification means using a traffic descriptor to categorize a packet within a specific group and to define that packet to make it accessible for quality of service (QoS) handling on the network. Using proper traffic classification, the network traffic is partitioned into multiple priority levels or classes of service. This lesson describes traffic classification best practices.

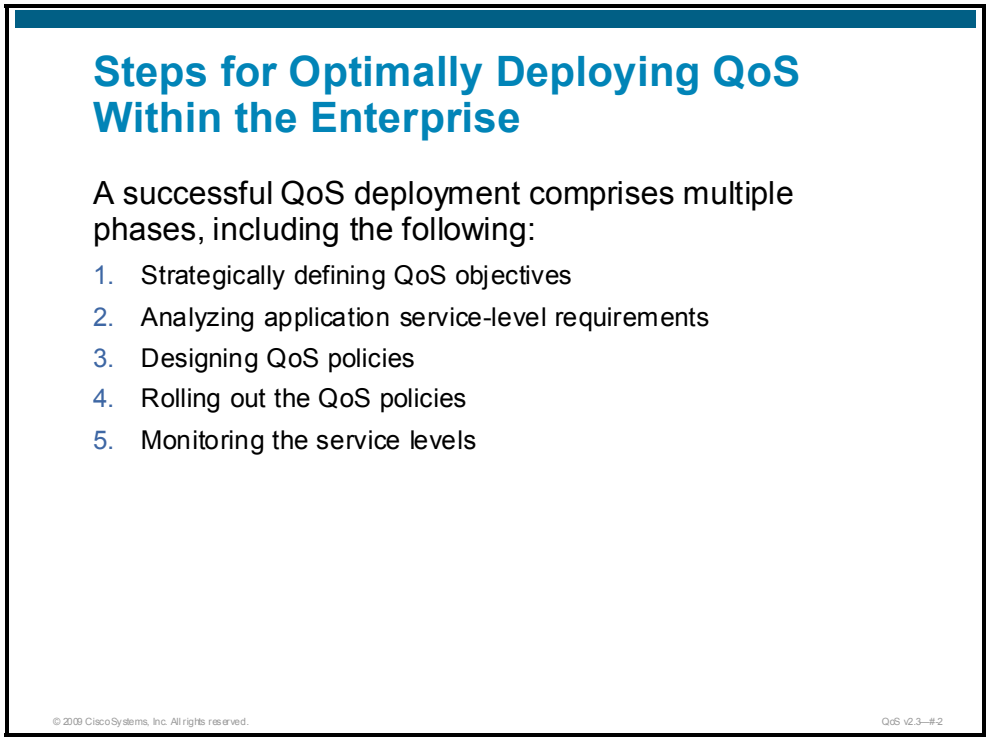
Objectives

Upon completing this lesson, you will be able to describe the set of classification practices that most closely represents Cisco QoS best practices. This ability includes being able to meet these objectives:

- List and describe the steps for optimally deploying QoS within an enterprise
- Explain how to begin a successful QoS deployment by strategically defining the business objectives to be achieved
- Explain the QoS requirements of the various application types
- Define some of the key QoS best-practices recommendations
- Explain how to map enterprise traffic classes into appropriate service provider traffic classes

Optimally Deploying QoS Within the Enterprise

This topic describes the steps required for optimally deploying QoS within an enterprise.



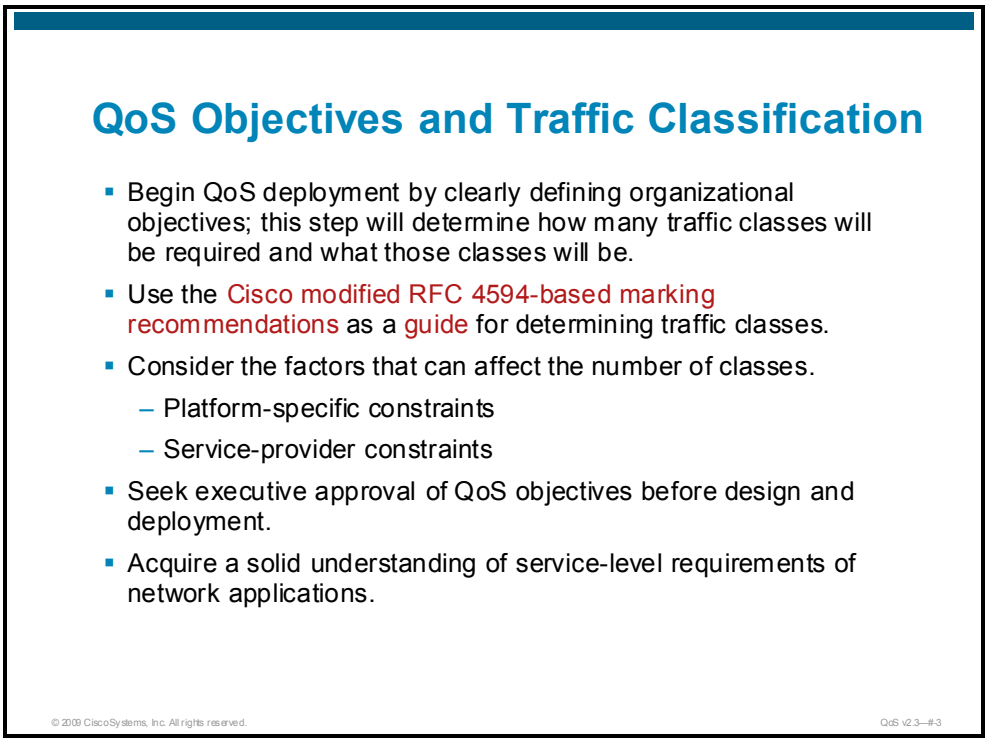
A successful QoS deployment comprises multiple phases, including the following:

1. Strategically defining the business objectives to be achieved via QoS.
2. Analyzing the service-level requirements of the various traffic classes to be provisioned for.
3. Designing and testing QoS policies prior to production-network rollout.
4. Rolling out the tested QoS designs to the production network.
5. Monitoring service levels to ensure that the QoS objectives are being met.

These phases may need to be repeated as business conditions change and evolve.

Strategically Defining QoS Objectives

This topic explains how to begin a successful QoS deployment by strategically defining the business objectives to be achieved and using the Cisco modified RFC 4594-based marking recommendations as a guide.



QoS technologies are the enablers for business and organizational objectives. Therefore, the way to begin a QoS deployment is not to activate QoS features simply because they exist, but to start by clearly defining the objectives of the organization. For example, one of the first questions that should arise during QoS planning is: “How many traffic classes should be provisioned for, and what should they be?”

To help answer these fundamental questions, organizational objectives such as the following must be defined:

- Is the objective to enable VoIP only, or is video also required?
- If video is required, is video-conferencing required or streaming video? Are both required?
- Are some applications considered mission-critical, and if so, which ones?
- Does the organization wish to suppress certain types of traffic, and if so, which types?

To help address these crucial questions and to simplify QoS, Cisco has adopted an initiative called the Cisco modified RFC 4594-based marking recommendations. The document is strategically designed to unify QoS within Cisco, from enterprise to service provider, and from engineering to marketing. It also provides uniform, standards-based recommendations to help ensure that QoS designs and deployments are unified and consistent.

Cisco Modified RFC 4594-Based Marking Recommendations

Application	L3 Classification		IETF
	PHB	DSCP	RFC
Network Control	CS6	48	RFC 2474
VoIP Telephony	EF	46	RFC 3246
Broadcast Video	CS5	40	RFC 2474
Multimedia Conferencing	AF41	34	RFC 2597
Real-Time Interactive/TelePresence	CS4	32	RFC 2474
Multimedia Streaming	AF31	26	RFC 2597
Call Signaling	CS3	24	RFC 2474
Low-Latency/Transactional Data	AF21	18	RFC 2597
OAM	CS2	16	RFC 2474
High-Throughput/Bulk Data	AF11	10	RFC 2597
Best Effort	DF	0	RFC 2474
Low-Priority/Scavenger Data	CS1	8	RFC 3662

Enterprises do not need to deploy all 12 classes of the modified RFC 4594-based marking recommendations model. This model is intended to be a forward-looking guide that considers as many classes of traffic with unique QoS requirements as possible. Familiarity with this model can assist in the smooth expansion of QoS policies to support additional applications as future requirements arise. However, at the time of QoS deployment, the enterprise needs to clearly define its organizational objectives, which will correspondingly determine how many traffic classes will be required.

This consideration should be tempered with the determination of how many application classes the networking administration team feels comfortable with deploying and supporting. Platform-specific constraints or service-provider constraints may also affect the number of classes of service.

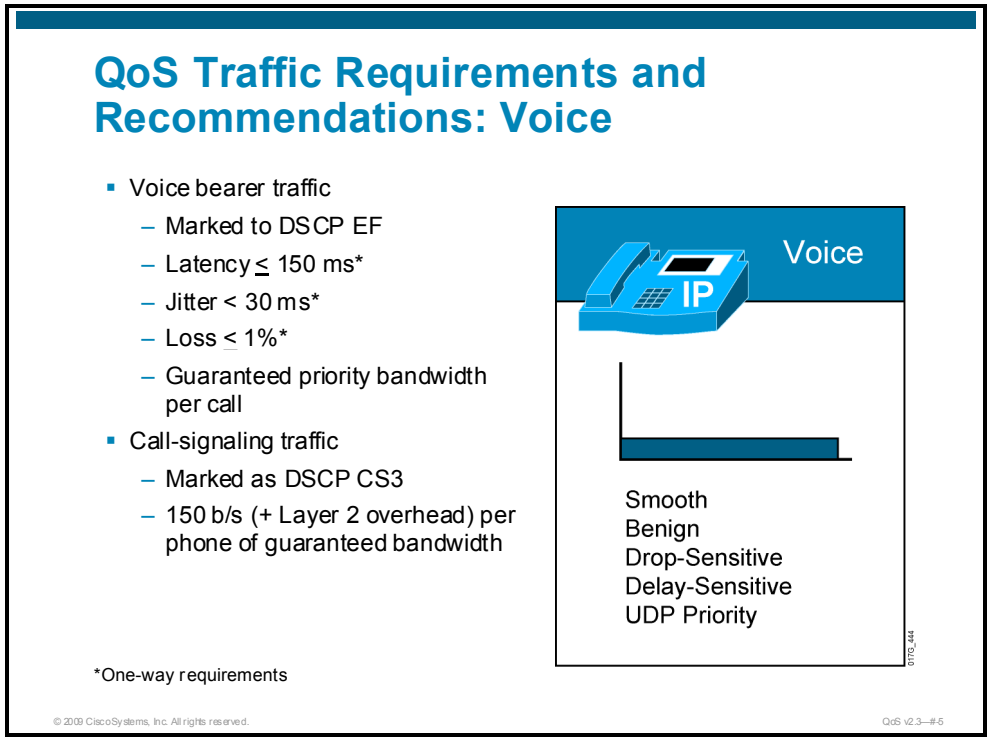
At this point you should also consider a migration strategy to allow the number of classes to be smoothly expanded as future needs arise, as shown in the figure.

Always seek executive endorsement of the QoS objectives prior to design and deployment. QoS is a system of managed unfairness and as such almost always bears political and organizational repercussions when implemented. To minimize the effects of these non-technical obstacles to deployment, address these political and organizational issues as early as possible, garnering executive endorsement whenever possible.

A strategic standards-based guide such as the Cisco modified RFC 4594-based marking recommendations coupled with a working knowledge of QoS tools and syntax is a prerequisite for any successful QoS deployment. However, you must also understand the service-level requirements of the various applications requiring preferential or deferential treatment within the network.

Analyzing Application Service-Level Requirements

This topic describes the QoS requirements of the various application types.



VoIP deployments require provisioning explicit priority servicing for VoIP (bearer stream) traffic and a guaranteed bandwidth service for call-signaling traffic. The key QoS requirements and recommendations for voice bearer traffic are as follows:

- Voice traffic should be marked to DSCP EF per the modified RFC 4594-based marking recommendations.
- Packet loss should be no more than 1 percent.
- One-way latency (mouth-to-ear) should be no more than 150 ms.
- Average one-way jitter should be targeted under 30 ms.
- Depending on the sampling rate, VoIP codec and Layer 2 media overhead, 21-320 kb/s of guaranteed priority bandwidth is required per call.

Voice quality is directly affected by all three QoS quality factors: loss, latency, and jitter.

- **Loss:** Loss causes voice clipping and skips. The industry-standard coded algorithms that are used in Cisco digital signal processors (DSPs) can correct for up to 30 ms of lost voice. For example, if a 20-ms sample of voice payload is used per VoIP packet, only a single voice packet can be lost during any given time. If two successive voice packets are lost, the 30-ms correctable window is exceeded and voice quality begins to degrade. VoIP networks are typically designed for very close to zero percent VoIP packet loss, with the only actual packet loss being due to L2 bit errors or network failures.

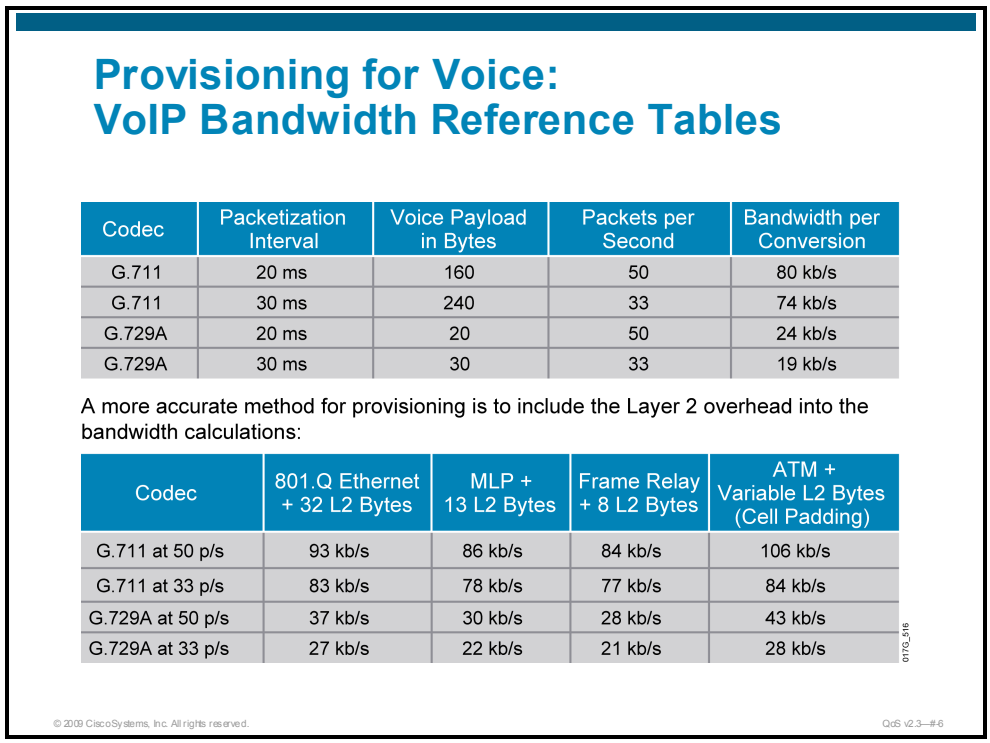
- **Latency:** Excessive latency can cause voice quality degradation. The goal commonly used in designing networks to support VoIP is the target specified by ITU standard G.114, which states that 150 ms of one-way, end-to-end (mouth-to-ear) delay ensures user satisfaction for telephony applications. A design should apportion this budget to the various components of network delay (propagation delay through the backbone, scheduling delay due to congestion, and the access link serialization delay) and service delay (due to VoIP gateway codec and de-jitter buffer). While the ITU G.114 states that a 150-ms, one-way delay budget is acceptable for high voice quality, lab testing has shown that there is a negligible difference in voice quality Mean Opinion Scores (MOS) using networks built with 200-ms delay budgets. Cisco thus recommends designing to the ITU standard of 150 ms, but if constraints exist where this delay target cannot be met, the delay boundary can be extended to 200 ms without significant impact on voice quality.
- **Jitter:** Jitter buffers, which are also known as play-out buffers, are used to change asynchronous packet arrivals into a synchronous stream by turning variable network delays into constant delays at the destination end systems. The role of the jitter buffer is to balance the delay and the probability of interrupted playout due to late packets. Late or out-of-order packets are discarded. If the jitter buffer is arbitrarily either large or small, it imposes unnecessary constraints on the characteristics of the network. A jitter buffer that is set too large adds to the end-to-end delay; this means that less delay budget is available for the network, so the network needs to support a delay target tighter than practically necessary. If a jitter buffer is too small to accommodate the network jitter, buffer underflows or overflows can occur. An underflow occurs when the buffer is empty when the codec needs to play out a sample. An overflow occurs when the jitter buffer is full and newly-arriving packets cannot be enqueued in the jitter buffer. Both jitter buffer underflows and overflows cause packets to be discarded. Adaptive jitter buffers aim to overcome these issues by dynamically tuning the jitter buffer size to the lowest acceptable value. Alternatively, the 30-ms value can be used as a jitter target; extensive lab testing has shown that when jitter consistently exceeds 30 ms, voice quality degrades significantly.

Because of its strict service-level requirements, VoIP is well suited to the Expedited Forwarding Per-Hop Behavior, as defined in RFC 3246 (formerly RFC 2598). It should therefore be marked to DSCP EF (46) and assigned strict priority servicing at each node, regardless of whether such servicing is done in hardware (as in Catalyst switches via hardware priority queuing) or in software (as in Cisco IOS routers via LLQ).

The following are key QoS requirements and recommendations for call-signaling traffic:

- Call-signaling traffic should be marked as DSCP CS3 per the modified RFC 4594-based marking recommendations. During migration, it may also be marked the legacy value of DSCP AF31.
- Voice control traffic requires 150 b/s (plus Layer 2 overhead) per phone of guaranteed bandwidth. More may be required, depending on the call signaling protocols in use.

Call-Signaling traffic was originally marked by Cisco IP Telephony equipment to DSCP AF31. However, the Assured Forwarding classes, as defined in RFC 2597, were intended for flows that could be subject to markdown and subsequently the aggressive dropping of marked-down values. Marking down and aggressively dropping call-signaling could result in noticeable delay-to-dial-tone (DDT) and lengthy call setup times, both of which generally translate to poor user experiences. The QoS Baseline (and subsequently, the modified RFC 4594-based marking recommendations) changed the marking recommendation for call-signaling traffic to DSCP CS3 because Class Selector code points, as defined in RFC 2474, were not subject to markdown and aggressive dropping. Each call signaling protocol has unique TCP/UDP ports and traffic patterns that should be taken into account when provisioning QoS policies for them.



The bandwidth consumed by VoIP streams (in b/s) is calculated by adding the VoIP sample payload (in bytes) to the 40-byte IP/UDP/RTP headers (assuming that cRTP is not in use), multiplying this value by the packetization rate (default of 50 packets per second) and then multiplying again by 8 to convert it to bits.

The first table in the figure details the bandwidth per VoIP flow at a default packet rate of 50 packets per second (p/s). This does not include Layer 2 overhead and does not take into account any possible compression schemes, such as cRTP.

Note	The Service Parameters menu in Cisco CallManager Administration can be used to adjust the packet rate. It is possible to configure the sampling rate above 30 ms, but this usually results in poor voice quality.

A more accurate method for provisioning VoIP is to include the Layer 2 overhead, which includes preambles, headers, flags, cyclic redundancy checks (CRCs), and ATM cell-padding. The amount of overhead per VoIP call depends on the Layer 2 technology used:

- 802.1Q Ethernet adds up to 32 bytes of Layer 2 overhead.
- Point-to-point protocol (PPP) adds 12 bytes of Layer 2 overhead.
- Multilink PPP (MLP) adds 13 bytes of Layer 2 overhead.
- Frame Relay adds 4 bytes of Layer 2 overhead; Frame Relay with FRF.12 adds 8 bytes.
- ATM adds varying amounts of overhead, depending on the cell padding requirements.

The second table in the figure shows a bandwidth provisioning example for voice that includes Layer 2 overhead. When determining the per-call bandwidth requirement for voice traffic, keep the following in mind:

- Codec type

- Packetization interval
- Layer 2 protocol overhead
- Bandwidth required for the voice control (signaling) traffic

Note A tool for quickly and accurately calculating VoIP bandwidth requirements can be found at the following site: http://tools.cisco.com/Support/VBC/jsp/Codec_Calc1.jsp.

Example: G.711 Voice Bearer Bandwidth Requirement Calculation

This example shows how to calculate the VoIP bearer bandwidth requirement for a single VoIP call using a G.711 codec (Layer 2 overhead not included):

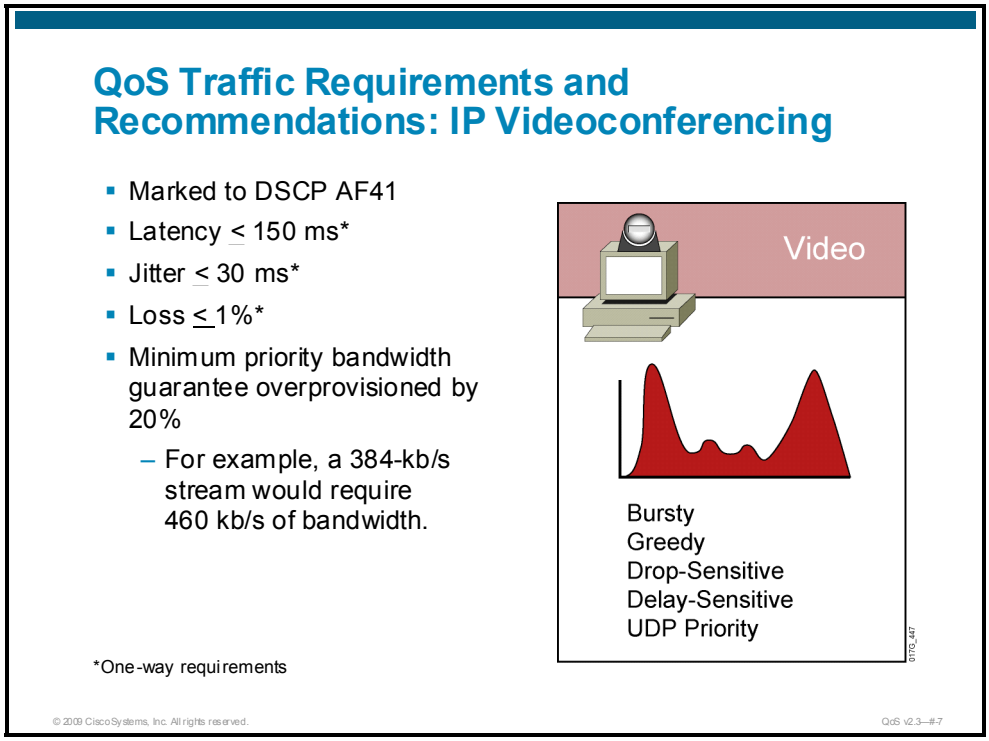
```
G.711              = 160 bytes payload size

Packet size        = payload size + IP/UDP/RTP headers
             = 160 bytes + 20 bytes + 8 bytes + 12 bytes
             = 200 bytes

Sampling Rate = 20 msec per sample = 50 samples per second

Bandwidth (bytes/sec) without Layer 2 overhead
= 200 bytes/packet x 50 packets/second
= 10000 bytes/second

Bandwidth (bits/sec) without Layer 2 overhead
= 10000 bytes/second  * 8 bits/byte
= 80000 bytes/second (80 kb/s)
```



When provisioning for IP Videoconferencing (IP/VC) traffic (referred to as Multimedia Conferencing in the modified RFC 4594-based marking recommendations), the following guidelines are recommended:

- IP/VC traffic should be marked to DSCP AF41; excess IP/VC traffic can be marked down by a policer to AF42 or AF43.
- Loss should be no more than one percent.
- One-way latency should be no more than 150 ms.
- Jitter should be no more than 30 ms.
- IP/VC queues should be overprovisioned by 20 percent to accommodate bursts.

Because IP/VC includes a G.711 audio codec for voice, it has the same loss, delay, and delay variation requirements as voice, but the traffic patterns of IP/VC are radically different from voice. For example, IP/VC traffic has varying packet sizes and extremely variable packet rates.

The IP/VC rate is the sampling rate of the video stream, not the actual bandwidth the video call requires. In other words, the data payload of IP/VC packets is filled with 384 kb/s worth of video and voice samples.

IP, UDP, and RTP headers (40 bytes per packet, uncompressed) need to be included in IP/VC bandwidth provisioning, as does the Layer 2 overhead of the media in use. Because (unlike VoIP) IP/VC packet sizes and rates vary, the header overhead percentage will vary as well, so an absolute value of overhead cannot be accurately calculated for all streams. Testing, however, has shown that a conservative rule of thumb for IP/VC bandwidth provisioning is to overprovision the priority bandwidth guarantee by 20 percent. For example, a 384-kb/s IP/VC stream would be adequately provisioned with an LLQ and CBWFQ of 460 kb/s.

Note The Cisco LLQ algorithm has been implemented to include a default burst parameter equivalent to 200 ms worth of traffic. Testing has shown that this burst parameter does not require additional tuning for a single IP/VC stream. For multiple streams, this burst parameter may be increased as required.

When addressing the QoS needs of streaming video traffic, the following guidelines are recommended:

- Streaming video (Multimedia Streaming), should be marked to AF31, as designated by the modified RFC 4594-based marking recommendations.
- Loss should be no more than 5 percent.
- Latency should be no more than 4 to 5 seconds, depending on video application buffering capabilities.
- There are no significant jitter requirements.
- Guaranteed bandwidth (CBWFQ) requirements depend on the encoding format and rate of the video stream.
- Non-organizational streaming video applications, such as entertainment videos, may be marked as scavenger (DSCP CS1) and assigned a minimal bandwidth (CBWFQ) percentage.

Streaming video applications may have more lenient QoS requirements for two reasons. First, they can be delay-insensitive. Second, they may be jitter-insensitive due to application buffering. However, streaming video may contain valuable content, such as e-learning applications or multicast company meetings, and therefore may require service guarantees.

Non-organizational video content (or video that is strictly entertainment-oriented in nature, such as music videos and humorous commercials) might be considered for a scavenger service. This means that these streams play if bandwidth exists, but they are the first to be dropped during periods of congestion.

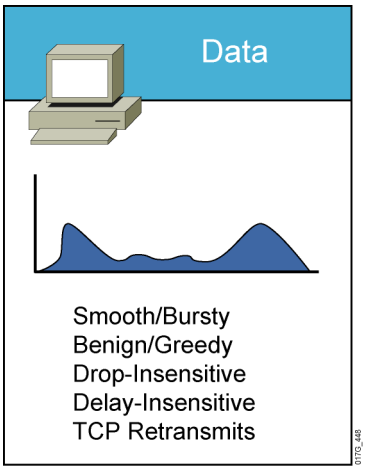
Example: Calculating the Bandwidth Requirement for a 384-kb/s Videoconference Stream

This example shows how to calculate the bandwidth requirement for a 384-kb/s IP/VC stream:

```
384 kb/s + (20% x 384 kb/s)    = 384 kb/s + 76.8 kb/s
                               = 460.8 kb/s
```

QoS Traffic Requirements: Data

- Best-Effort Data traffic
  - Marked to DSCP DF
  - At least 25% of bandwidth
- Bulk (High-Throughput) Data traffic
  - Marked to DSCP AF11
  - Moderate bandwidth guarantee
- Transactional (Low-Latency) Data traffic
  - Marked to DSCP AF21
  - Adequate bandwidth guarantee for interactive, foreground operations
- Scavenger (Low-Priority) Data traffic
  - Marked to DSCP CS1
  - Minimal bandwidth queue that is the first to starve during congestion



There are hundreds of thousands of data networking applications. Some are TCP, others are UDP; some are delay sensitive, others are not; some are bursty in nature, others are steady; some are lightweight, others require high bandwidth. Not only do applications vary one from another, but the same application can vary significantly from one version to another. The modified RFC 4594-based marking recommendations identify four main classes of data traffic, according to their general networking characteristics and requirements. These classes are best effort, high-throughput data, low-latency data, and low-priority data.

The best-effort class is the default class for all data traffic. An application is removed from the default class only if it is selected for preferential or deferential treatment. For best-effort data traffic, Cisco recommends the following guidelines:

- Best-effort traffic should be marked to default forwarding (DF) or DSCP 0.
- Adequate bandwidth should be assigned to the best-effort class as a whole, because the majority of applications default to this class. Traffic in this class should be provisioned with a dedicated queue. It is also recommended that you enable WRED on this class. However, because all the traffic in this class is marked to DSCP 0, the congestion avoidance mechanism is essentially random early detection (RED).

The bulk (high-throughput) data class is intended for applications that are relatively non-interactive and drop-insensitive and that typically span their operations over a long period of time as background occurrences. Such applications include FTP, backup operations, database synchronizing or replicating operations, content distribution, and any other type of background operation. Because most background applications are TCP-based file transfers, these applications, if left unchecked, can take excessive network resources away from more interactive, foreground applications.

For bulk data traffic, Cisco recommends the following guidelines:

- Bulk data traffic should be marked to DSCP AF11; excess bulk data traffic can be marked down by a policer to AF12; violating bulk data traffic may be marked down further to AF13 or dropped.
- Traffic in this class should be provisioned with a moderate bandwidth guarantee.

The transactional (low-latency) data class is a combination of two similar types of applications: transactional data client-server applications and interactive messaging applications. The response time requirement separates transactional data client-server applications from generic client-server applications. For example, with transactional data client-server applications such as SAP, PeopleSoft, and Data-Link Switching (DLSw+), the transaction is a foreground operation; the user waits for the operation to complete before proceeding. Email is not considered a transactional data client-server application, because most email operations occur in the background and users do not usually notice even several-hundred-millisecond delays in mailspool operations.

For transactional data traffic, Cisco recommends the following guidelines:

- Transactional data traffic should be marked to DSCP AF21; excess transactional data traffic can be marked down by a policer to AF22; violating transactional data traffic can be marked down further to AF23 or dropped.
- Transactional data traffic should have an adequate bandwidth guarantee for the interactive, foreground operations that they support.

The scavenger (low-priority) data class is intended for non-business related traffic flows, such as data or media applications that are entertainment-oriented. The approach of a less-than-best-effort service class for non-business applications (as opposed to shutting them down entirely) has proven to be a popular, political compromise. These applications are permitted on enterprise networks as long as resources are always available for business-critical applications. However, as soon the network experiences congestion, this class is the first to be penalized and is aggressively dropped. Furthermore, the scavenger class can be utilized as part of an effective strategy for DoS and worm attack mitigation. Examples of scavenger traffic include YouTube, Xbox 360 movies, iTunes, and BitTorrent.

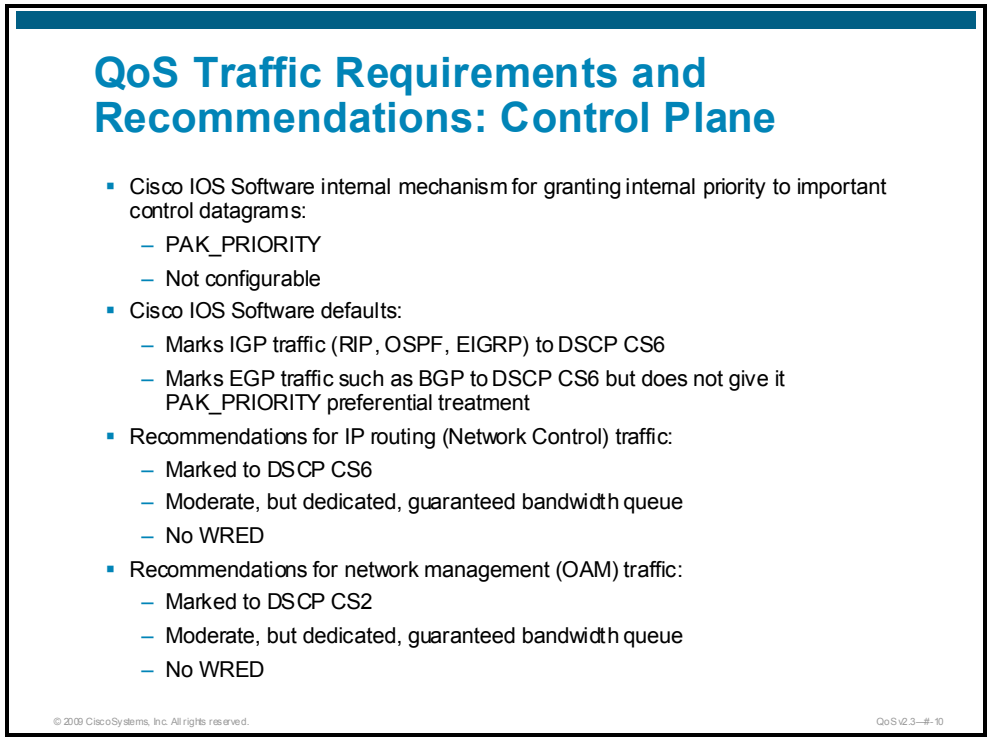
For scavenger data traffic, the following guidelines are recommended:

- Scavenger data traffic should be marked CS1.
- Scavenger data traffic should be provisioned with a minimal bandwidth queue that is the first to starve if network congestion occurs.

Grouping Data Applications into Classes

Application Class	Example Applications	Applications/ Traffic Properties	Packet/ Message Sizes
Transactional	PeopleSoft (Vantive), Microsoft SQL Server	Typically use a client-server protocol model. User initiated client-based queries followed by server response. Query response may consist of many messages between client and server or of many TCP and FTP sessions running simultaneously.	Depends on application; could be anywhere from 1 KB to 50 MB
Bulk	Network-based backups, Microsoft Outlook	Long file transfers. Always invokes TCP congestion management.	Average message size 64 KB or greater
Best Effort	All non-critical traffic, HTTP web browsing, other miscellaneous traffic	Varies	Varies

When classifying the different network applications into traffic classes, try to group applications with common characteristics and QoS requirements together into the same traffic class. The figure shows some applications and the generic networking characteristics that determine the data application class for which they are best suited.



Unless the network is up and running, QoS is irrelevant. Therefore, it is critical to provision QoS for control plane traffic, which includes IP routing traffic and network management.

By default, Cisco IOS Software (in accordance with RFC 791 and RFC 2474) marks interior gateway protocol (IGP) traffic such as Routing Information Protocol (RIP/RIPv2), Open Shortest Path First (OSPF), and Enhanced Interior Gateway Routing Protocol (EIGRP) to DSCP CS6. However, Cisco IOS Software also has an internal mechanism for granting internal priority to important control datagrams as they are processed within the router. This mechanism is called PAK_PRIORITY.

As datagrams are processed though the router and down to the interfaces, they are internally encapsulated with a small packet header, referred to as the PAKTYPE structure. Within the fields of this internal header there is a PAK_PRIORITY flag that indicates the relative importance of control packets to the internal processing systems of the router. PAK_PRIORITY designation is a critical internal Cisco IOS Software operation and, as such, is not administratively configurable in any way.

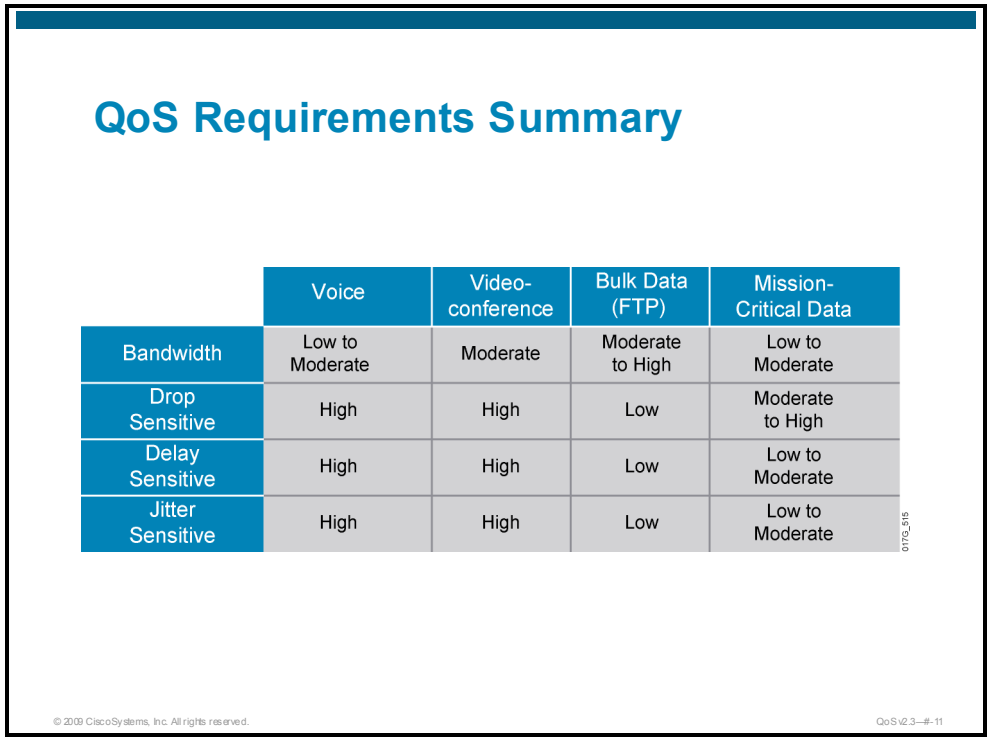
Exterior gateway protocol (EGP) traffic such as Border Gateway Protocol (BGP) traffic is marked by default to DSCP CS6, but does not receive such PAK_PRIORITY preferential treatment and may need to be explicitly protected in order to maintain peering sessions.

When addressing the QoS needs of IP routing (network control) traffic, Cisco recommends the following guidelines:

- IP routing traffic should be marked to DSCP CS6; this is default behavior on Cisco IOS platforms. Cisco IOS Software automatically marks IP routing traffic to DSCP CS6.
- IP routing traffic should be provisioned with a moderate, but dedicated, guaranteed bandwidth queue.
- WRED should not be enabled on this class, because network control traffic should not be dropped. If the network control class is experiencing drops, the bandwidth allocated to it should be re-provisioned.

When addressing the QoS needs of network management (OAM) traffic, Cisco recommends the following guidelines:

- Network management traffic should be marked to DSCP CS2.
- Network management traffic should be provisioned with a moderate, but dedicated, guaranteed bandwidth queue.
- WRED should not be enabled on this class. If the OAM class is experiencing drops, the bandwidth allocated to it should be re-provisioned.



The table summarizes the key QoS requirements (bandwidth, packet loss, delay, and jitter) for some of the major categories of applications.

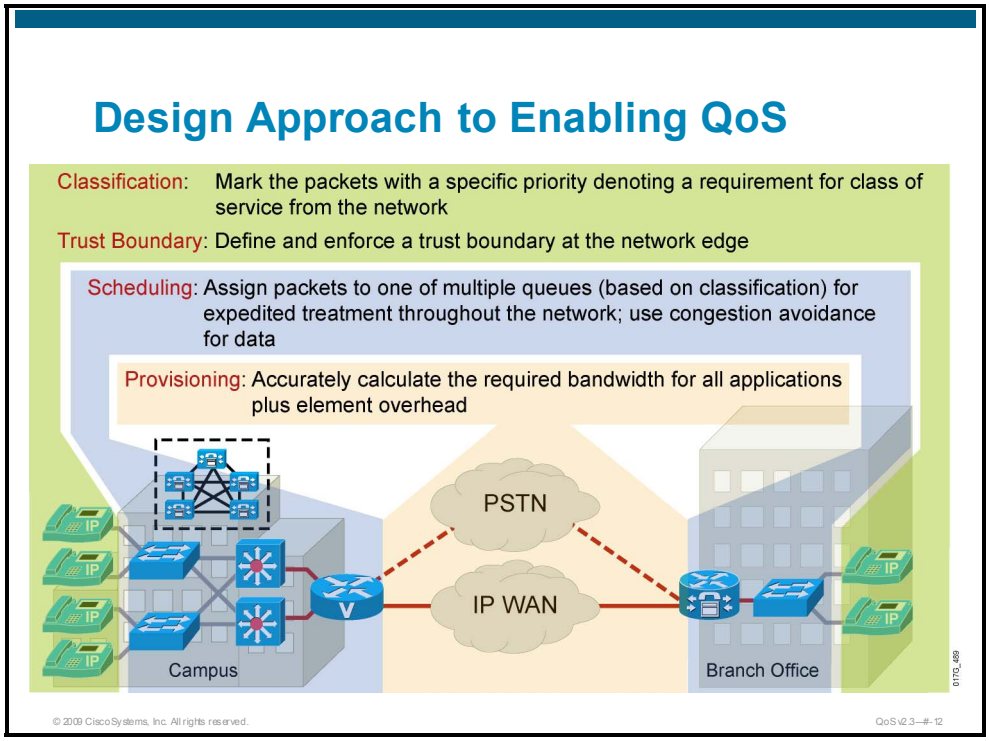
Example: QoS Requirements of the Major Applications Category

VoIP and IP/VC applications share the same low-drop, low-delay, and low-jitter requirements. For example, a latency of less than 150 ms, a jitter of less than 30 ms, and a packet loss of less than one percent are typically required.

Bulk applications such as FTP are less sensitive to drop, delay, and jitter, but generally require more bandwidth than real-time traffic such as voice.

Designing the QoS Policies

This topic describes best-practice QoS design principles.

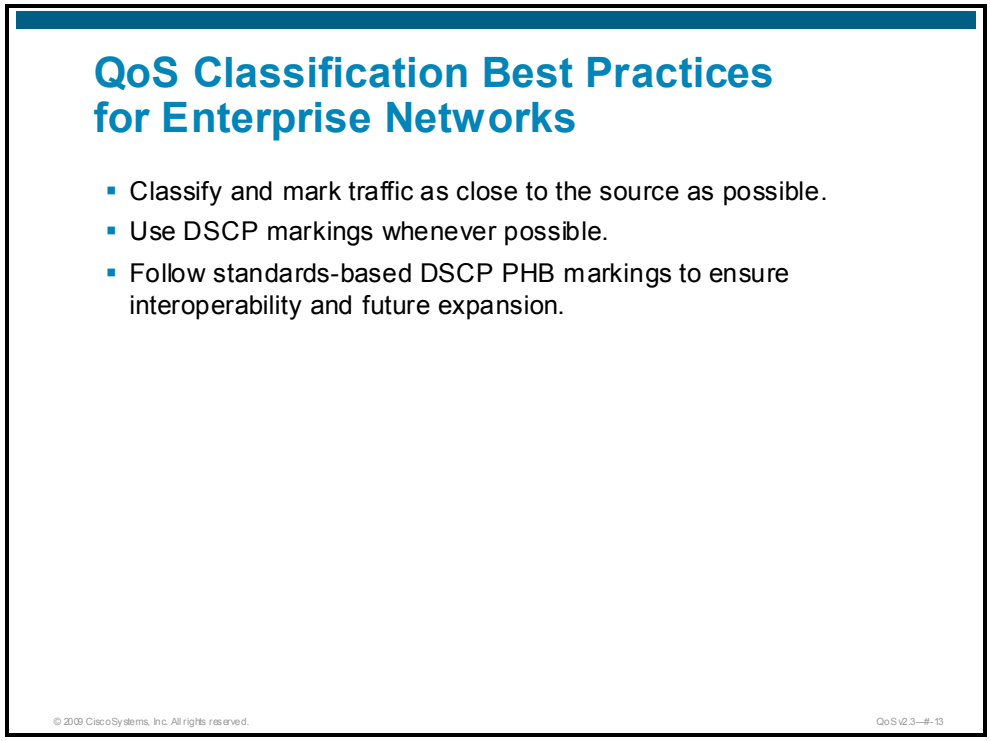


Once a QoS strategy has been defined and the application requirements are understood, end-to-end QoS policies can be designed for each device and interface, as determined by its role in the network infrastructure. Because the Cisco QoS toolset provides many QoS design and deployment options, a few succinct design principles can help simplify strategic QoS deployments.

For example, one such design principle is to always enable QoS policies in hardware, rather than software, whenever a choice exists. Cisco IOS routers perform QoS in software, which places incremental loads on the CPU, depending on the complexity and functionality of the policy. Cisco Catalyst switches, on the other hand, perform QoS in dedicated hardware ASICS and therefore do not tax their main CPUs to administer QoS policies. This allows complex policies to be applied at line rates at even gigabit or ten-gigabit speeds.

Other simplifying best-practice QoS design principles include the following:

- Classification and marking principles
- Policing and markdown principles
- Queueing and dropping principles



The first element of a QoS policy is to identify the traffic to be treated differently. By classifying the application traffic into different traffic classes, a baseline methodology is set to provide end-to-end QoS. DiffServ enables this classification by using the differentiated services code point (DSCP) field. Using DiffServ, a properly designed network can deliver assured bandwidth, low latency, low jitter, and low packet loss for voice while simultaneously ensuring slices of available bandwidth to other traffic classes. Packets entering a DiffServ domain (a collection of DiffServ routers) can be classified in a variety of ways, as follows:

- IP source and destination addresses
- Layer 4 protocol and port numbers
- Incoming interface
- MAC address
- IP Precedence
- DSCP value
- Layer 2 information (Frame Relay discard eligible [DE] bits, Ethernet 802.1p bits)
- NBAR, the Cisco value-added mechanism

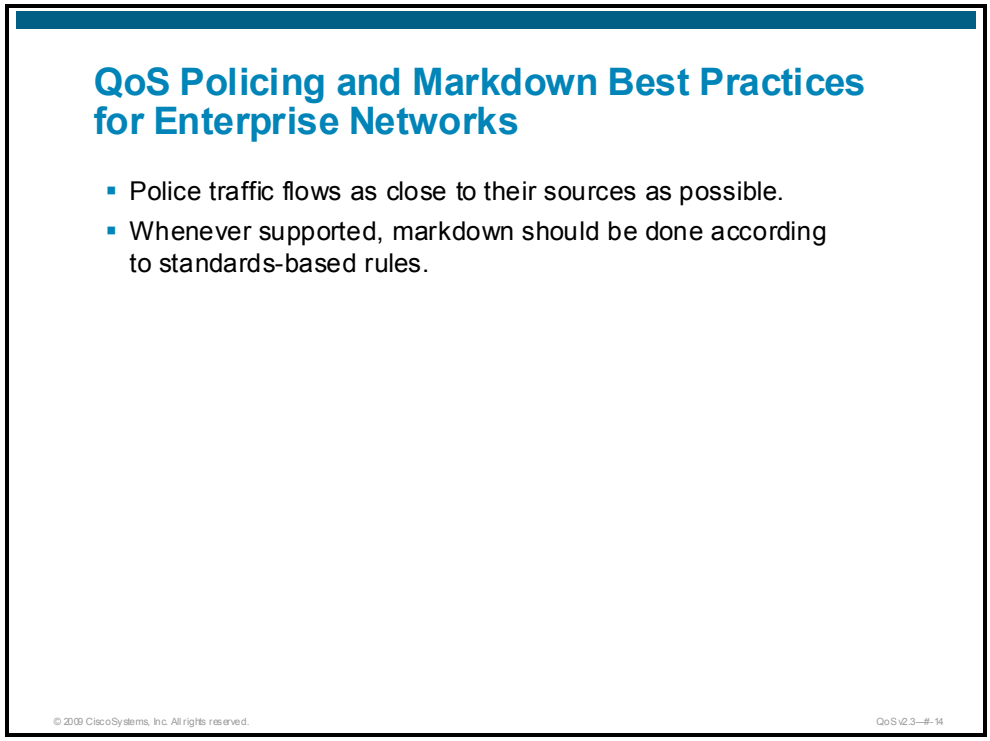
It is best practice to classify and mark applications as close to their sources as technically and administratively feasible. This principle promotes end-to-end Differentiated Services and Per-Hop Behaviors (PHBs). Do not trust markings that can be set by users on their PCs or other similar devices, because users can easily abuse provisioned QoS policies if permitted to mark their own traffic. For example, if DSCP EF receives priority services throughout the enterprise, a PC can be easily configured to mark all the traffic of the user to DSCP EF, thus hijacking network priority queues to service non-real-time traffic. Such abuse could easily ruin the service quality of real-time applications like VoIP throughout the enterprise.

Following this rule, it is further recommended to use DSCP markings whenever possible, because these are end-to-end, more granular and more extensible than Layer 2 markings. Layer 2 markings are lost when media changes (such as a LAN-to-WAN/VPN edge). There is also less marking granularity at Layer 2. For example, 802.1Q/p CoS supports only 3 bits (values 0 to 7), as does MPLS EXP. Therefore, only up to eight classes of traffic can be supported at Layer 2, and inter-class relative priority (such as RFC 2597 Assured Forwarding Drop Preference markdown) is not supported. On the other hand, Layer 3 DSCP markings allow for up to 64 classes of traffic, which is more than enough for most enterprise requirements for the foreseeable future.

As the line between enterprises and service providers continues to blur and the need for interoperability and complementary QoS markings is critical, you should follow standards-based DSCP PHB markings to ensure interoperability and future expansion. Because the modified RFC 4594-based marking recommendations are standards-based, enterprises can easily adopt these markings to interface with service provider classes of service. Network mergers are also easier to manage when you use standards-based DSCP markings.

In an enterprise environment, the QoS policies should allow critical business applications to receive requisite resources, while ensuring that other applications are not neglected. QoS policies should also ensure the quality of real-time traffic, such as voice and video. QoS policies may also need to prevent non-business-related network traffic (scavenger traffic), such as file-sharing, from taking up too much of the network bandwidth.

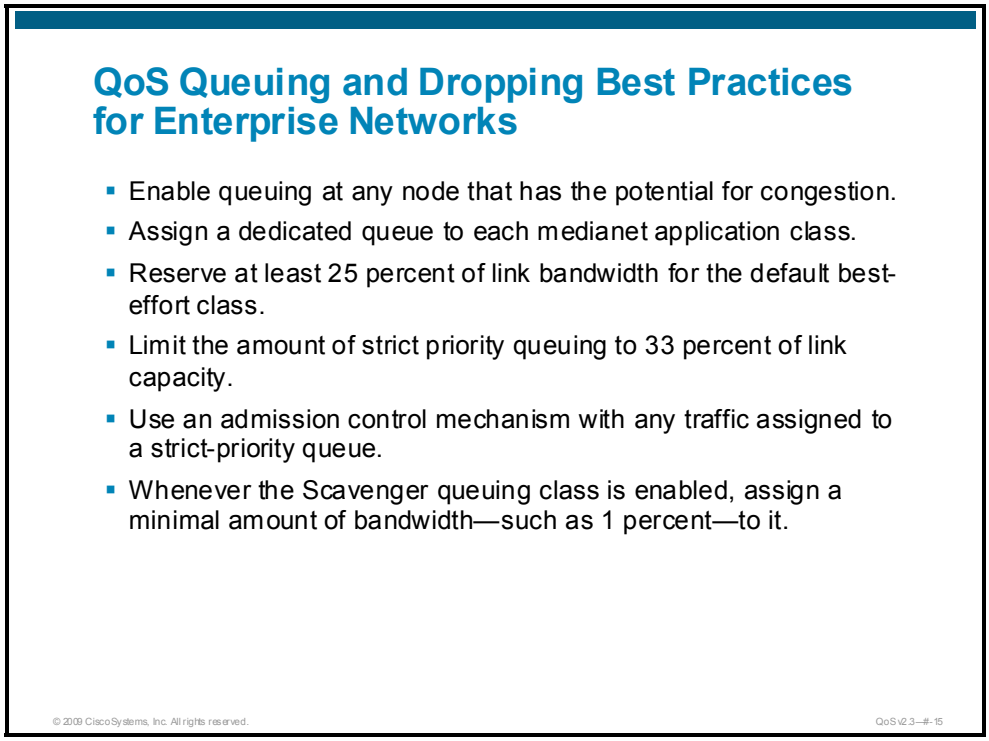
Network administrators often cannot justify continual upgrade of the link speeds in their networks. Cisco IOS Software QoS features provide an alternative solution to link upgrade by managing the links efficiently to meet the application demands. Use QoS mechanisms such as Multilink PPP link fragmentation and interleaving (MLP LFI), compressed Real-Time Transport Protocol (cRTP), class-based weighted fair queuing (CBWFQ), and low-latency queuing (LLQ) to allow the most efficient distribution of the available bandwidth among the applications.



There is little reason to forward unwanted traffic only to police and drop it at a subsequent node, especially when the unwanted traffic is the result of DoS or worm attacks. The overwhelming volume of traffic that such attacks can create can cause network outages by driving network device processors to their maximum levels. Therefore, you should police traffic flows as close to their sources as possible. This principle applies also to legitimate flows. DoS and worm-generated traffic can masquerade under legitimate, well-known TCP and UDP ports and cause extreme amounts of traffic to be poured onto the network infrastructure. Such excesses should be monitored at the source and marked down appropriately.

Whenever supported, markdown should be done according to standards-based rules, such as RFC 2597 (Assured Forwarding PHB Group). For example, excess traffic marked to AFx1 should be marked down to AFx2 (or AFx3, whenever dual-rate policing such as that defined in RFC 2698 is supported). Following such markdowns, congestion management policies, such as DSCP-based WRED, should be configured to drop AFx3 more aggressively than AFx2, which in turn should be dropped more aggressively than AFx1.

However, Cisco Catalyst switches do not currently perform DSCP-based WRED, so this standards-based strategy cannot be implemented fully at this time. As an alternative workaround, single-rate policers can be configured to mark down excess traffic to DSCP CS1 (scavenger); dual-rate policers can be configured to mark down excess traffic to AFx2, while marking down violating traffic to DSCP CS1. Traffic marked as scavenger would then be assigned to a “less-than-best-effort” queue. Such workarounds yield an overall effect similar to the standards-based policing model. However, when DSCP-based WRED is supported on all routing and switching platforms, you should mark down assured forwarding classes by RFC 2597 rules to comply more closely with this standard.



Critical applications such as VoIP require service guarantees regardless of network conditions. The only way to provide service guarantees is to enable queuing at any node that has the potential for congestion, regardless of how rarely this may occur. This principle applies not only to campus-to-WAN or VPN edges, where speed mismatches are most pronounced, but also to campus access-to-distribution or distribution-to-core links, where oversubscription ratios create the potential for congestion. There is simply no other way to guarantee service levels than by enabling queuing wherever a speed mismatch exists.

Additionally, because each medianet application class has unique service-level requirements, each should optimally be assigned a dedicated queue. However, on platforms bounded by a limited number of hardware or service provider queues, no fewer than four queues would be required to support medianet QoS policies, specifically:

- Real-time queue (to support an RFC 3246 EF PHB service)
- Guaranteed-bandwidth queue (to support RFC 2597 AF PHB services)
- Default queue (to support an RFC 2474 DF service)
- Bandwidth-constrained queue (to support an RFC 3662 scavenger service)

When provisioning queuing, some best-practice principles also apply. For example, as discussed previously, the best-effort class is the default class for all data traffic. Only if an application has been selected for preferential or deferential treatment is it removed from the default class. Because many enterprises have several hundred data applications—if not thousands—running over their networks, you must provision adequate bandwidth for this class as a whole to handle the sheer volume of applications that default to it. Therefore, it is recommended that you reserve at least 25 percent of link bandwidth for the default best-effort class.

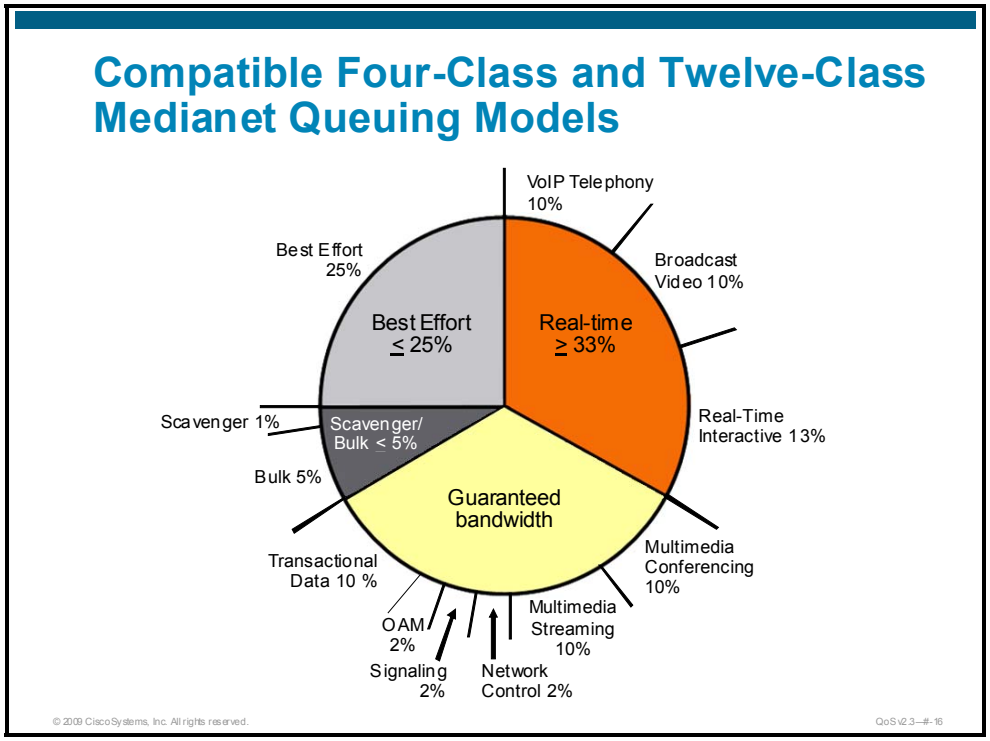
Not only does the best-effort class of traffic require special bandwidth provisioning consideration, so does the highest class of traffic, sometimes referred to as the “real-time” or “strict priority” class (which corresponds to RFC 3246: An Expedited Forwarding Per-Hop Behavior). The amount of bandwidth assigned to the real-time queuing class is variable. However, if you assign too much traffic for strict priority queuing, the overall effect is a dampening of QoS functionality for non-real-time applications. The goal of convergence is to enable voice, video, and data to transparently co-exist on a single network. When real-time applications such as voice or interactive video dominate a link (especially a WAN or VPN link), data applications will fluctuate significantly in their response times, destroying the transparency of the converged network.

Cisco Technical Marketing testing has shown a significant decrease in data application response times when real-time traffic exceeds one-third of link bandwidth capacity. Extensive testing and customer deployments have shown that a general best-queuing practice is to limit the amount of strict priority queuing to 33 percent of link capacity. This strict priority queuing rule is a conservative and safe design ratio for merging real-time applications with data applications.

Cisco IOS Software allows the abstraction (and thus configuration) of multiple strict-priority LLQs. In such a multiple LLQ context, this design principle would mean that the sum of all LLQs should be within one-third of link capacity.

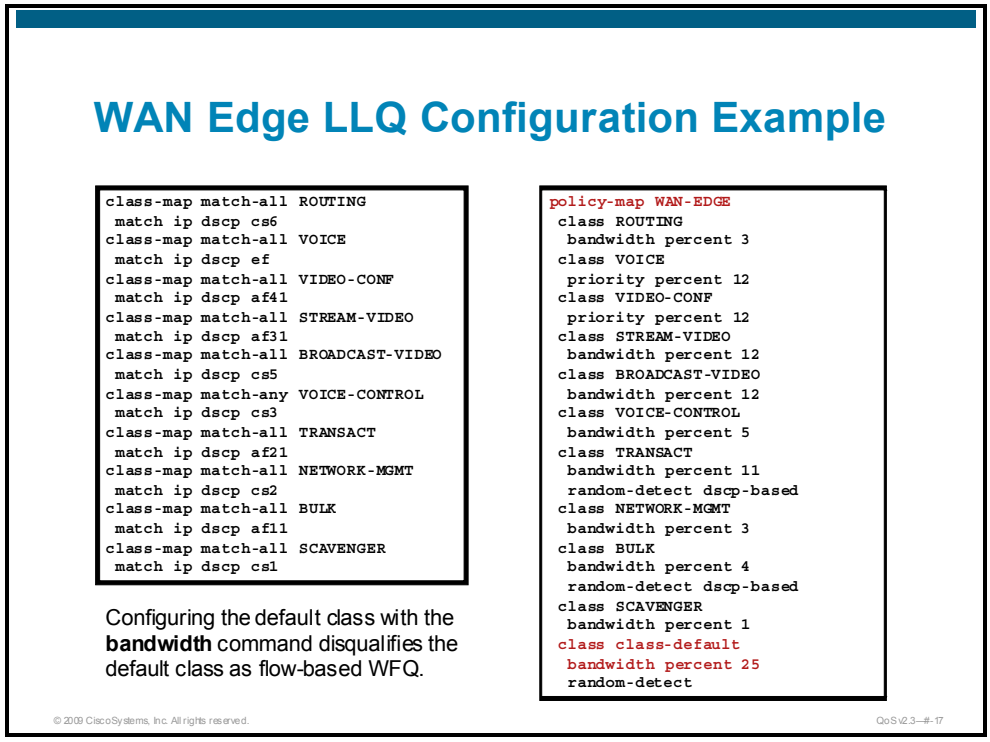
This strict priority queuing rule (limit to 33 percent) is simply a best-practice design recommendation and is not a mandate. There may be cases where specific business objectives cannot be met while holding to this recommendation. In such cases, enterprises must provision according to their detailed requirements and constraints. However, it is important to recognize the tradeoffs involved with over-provisioning strict priority traffic and its negative performance impact on non-real-time-application response times.

Whenever a scavenger queuing class is enabled, it should be assigned a minimal amount of bandwidth. On some platforms, queuing distinctions between bulk data and scavenger traffic flows cannot be made because queuing assignments are determined by CoS values and these applications share the same CoS value of 1. In such cases, you can assign the scavenger and bulk queuing class a bandwidth percentage of 5. If you can uniquely assign scavenger and bulk data to different queues, you should assign the scavenger queue a bandwidth percentage of 1.



The queuing rules presented in this topic are summarized in the figure, where the inner pie chart represents a hardware or service provider queuing model that is limited to four queues, and the outer percentages represent a corresponding, more granular queuing model that is not bound by such constraints.

Example: LLQ Example on the Enterprise WAN Edge Router



The figure illustrates an example of an enterprise WAN-edge router configuration using LLQ with class-based WRED on certain data traffic classes. In this example, the maximum reservable bandwidth on the link is set to 100 percent, so that up to 100 percent of the link bandwidth can be guaranteed among the various classes.

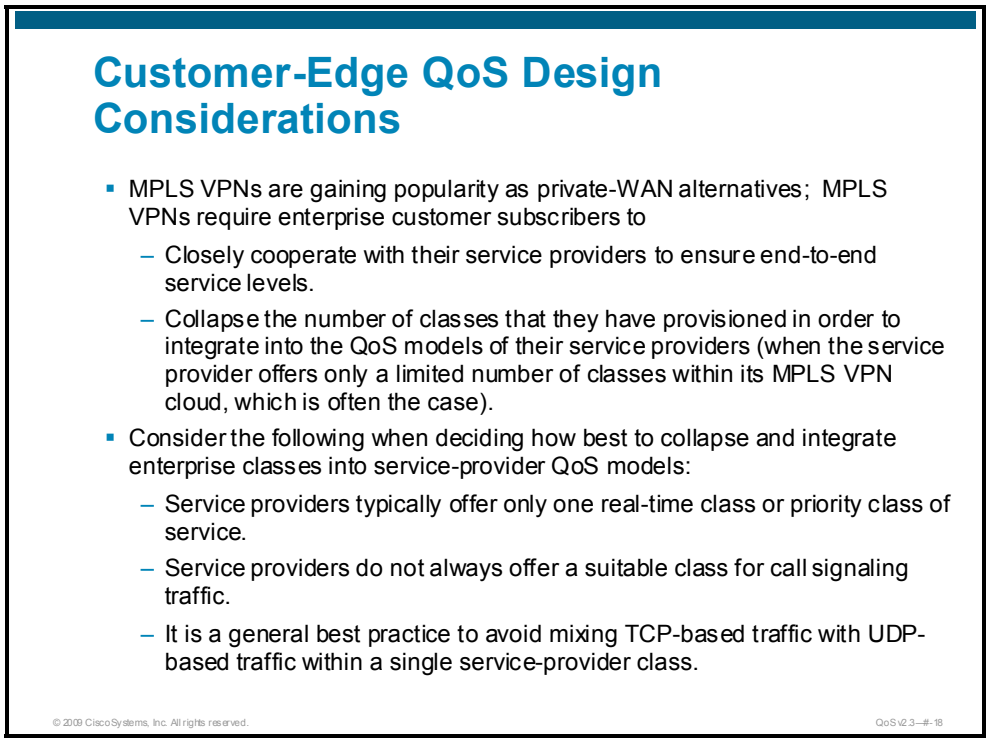
The example in the figure assumes that the markings for the different traffic classes are already done at the access or distribution layer within the campus network.

Currently (except for the Cisco 7500 Series routers) all traffic classes except for the default traffic class support only FIFO queuing within the class. The default traffic class can support either FIFO or WFQ within the class. However, if the default traffic class is allocated a minimum bandwidth, WFQ is not supported in the default traffic class.

Note	In Cisco IOS Software Release 12.4, the **match ip dscp** command is replaced by the **match dscp** command.

Enterprise-to-Service Provider QoS Class Mapping

This topic describes how to map enterprise traffic classes into traffic classes that are appropriate for service-provider use.



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 the policies of their service provider.

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 the QoS models of their service provider. The following caveats should be considered 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 IP/VC (each of which should be provisioned with strict priority treatment) over the 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 (less than 768 kb/s). 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.

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 that call signaling be protected. Service providers might not always offer a suitable class for call-signaling traffic itself. Therefore, the enterprise must determine which other traffic classes to mix with call signaling.

On links where serialization is not an issue, 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 application within the class.

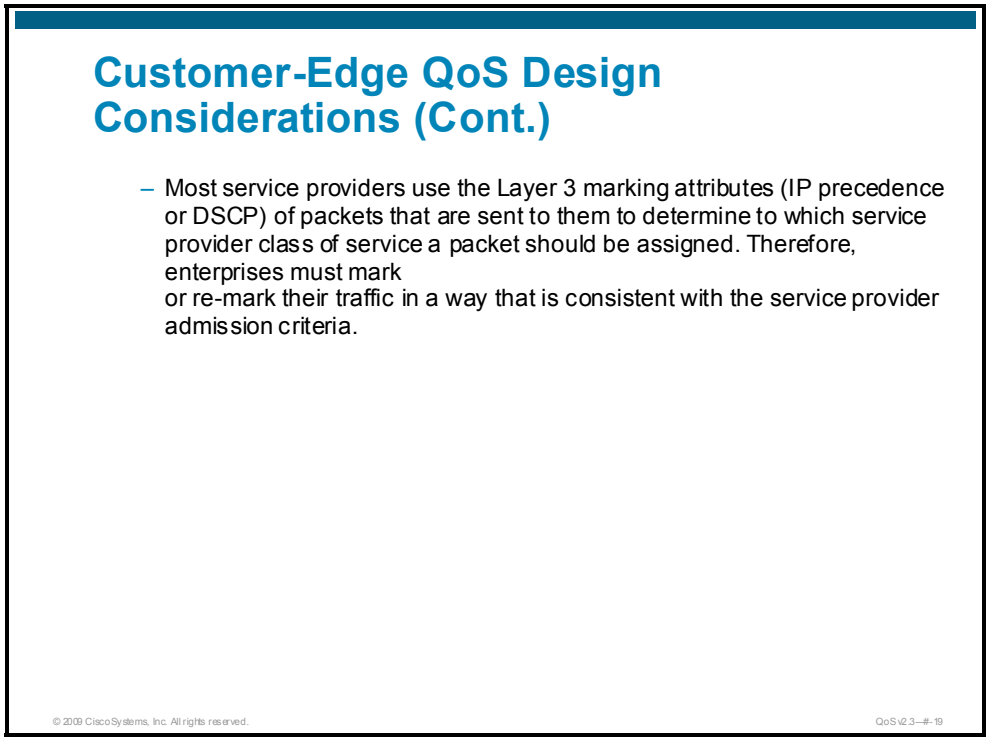
Mixing TCP with UDP

It is a general best practice to avoid mixing 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 and UDP dominance.

TCP starvation and 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 is observed because WRED (for the most part) manages congestion only on TCP-based flows.

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 the Layer 3 marking attributes (IP precedence or DSCP) of packets that are sent to them to determine to which service provider class of service a packet should be assigned. Therefore, enterprises must mark or re-mark their traffic in a way that is consistent with the service-provider admission criteria. Additionally, service providers might re-mark at Layer 3 out-of-contract traffic within their cloud; this can 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 egress edge of the customer edge (CE), rather than within the campus. This is because service-provider service offerings are likely to 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 must be marked to the same DiffServ code point value to gain admission to the appropriate queue. For example, on high-speed links, you might want to send voice, IP/VC, and call signaling to the service provider real-time class. If the service-provider class admits only DSCP EF and CS5, two of these applications would have to share a common code point. The following example, in which IP/VC and call signaling are re-marked to share DSCP CS5, shows how this can be done.

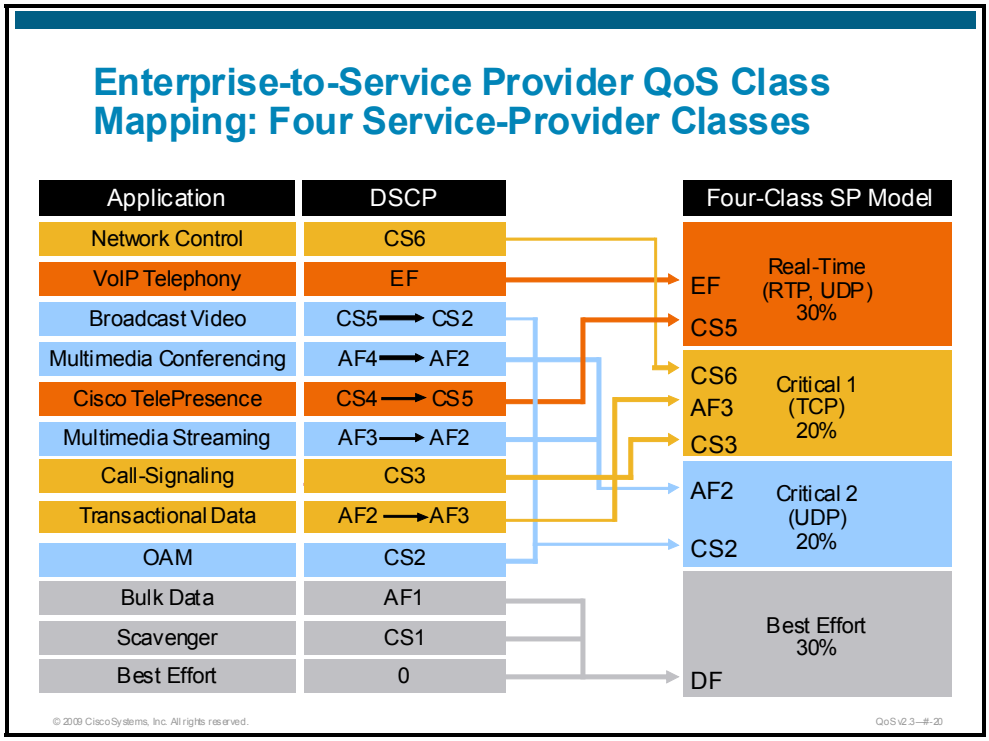
```
class-map match-any VOIP-TELEPHONY
 match ip dscp ef
class-map match-all MULTIMEDIA-CONFERENCING
 match ip dscp af41
class-map match-any CALL-SIGNALING
 match ip dscp cs3
!
policy-map CE-EGRESS-EDGE
 class VOIP-TELEPHONY
```

```
  priority percent 18
 class MULTIMEDIA-CONFERENCING
  priority percent 15
  set ip dscp cs5            ! Multimedia Conferencing 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
```

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, it might present minor difficulties to enterprises that require consistent end-to-end Layer 3 marking. In such cases, the enterprise can choose to apply re-marking policies as traffic is received back from the service provider MPLS VPN (on the ingress direction of the enterprise CE). Class-based marking can be used again because it supports not only access lists for classification, but also NBAR.

Continuing and expanding on the previous example, the enterprise wants to restore the original markings that it set for IP/VC and call signaling. Additionally, it wants to restore original markings for Oracle traffic (which it originally marked AF21 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. The following example 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.

```
class-map match-all REMARKED-MULTIMEDIA-CONFERENCING
  match ip dscp cs5              ! Interactive-Video must be CS5 AND UDP
  match access-group 101
!
class-map match-all REMARKED-CALL-SIGNALING
  match ip dscp cs5              ! Call-Signaling must be CS5 AND TCP
  match access-group 102
!
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-MULTIMEDIA-CONFERENCING
  set ip dscp af41               ! Restores Interactive-Video marking to AF41
 class REMARKED-CALL-SIGNALING
  set ip dscp cs3                ! Restores Call-Signaling marking to CS3
 class REMARKED-ORACLE
  set ip dscp af21               ! Restores Oracle marking to AF21
 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
```



In the model shown in the figure, the service provider offers four classes of service. Because there are so few classes to choose from in this example, Cisco TelePresence may need to be combined with another application. It is highly recommended not to combine Cisco TelePresence with any unbounded application (an application without admission control) within a single service provider class, because doing so could lead to class congestion and result in Cisco TelePresence drops (with or without WRED enabled on the service provider class). Therefore, there are two options in such a design:

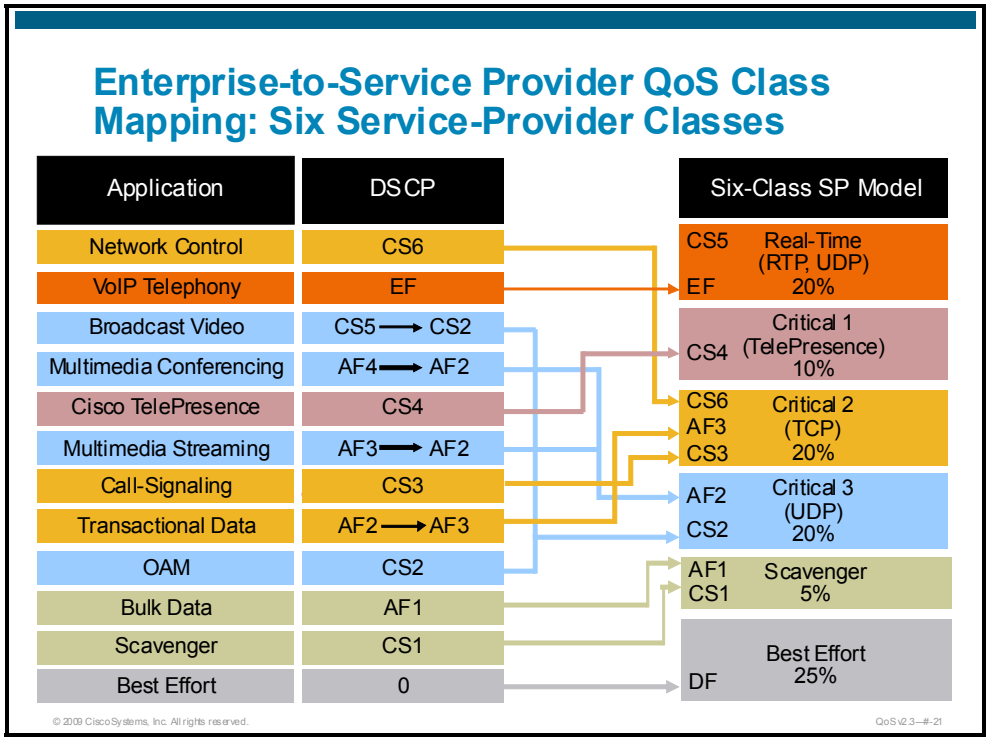
- Assign Cisco TelePresence to the service provider real-time class along with voice.
- Assign Cisco TelePresence to a dedicated non-priority service-provider class.

In this example, Cisco TelePresence is assigned to the service provider real-time class.

In the four-class service provider model, there is a real-time class, a default best-effort class, and two additional non-priority traffic classes. In this case, the enterprise administrator may elect to separate TCP-based applications from UDP-based applications by using these two non-priority service provider traffic classes. Specifically, if voice and Cisco TelePresence are the only applications to be assigned to the service provider real-time class, broadcast video, multimedia conferencing, multimedia streaming, and OAM traffic (which is largely UDP-based) can all be assigned to the service provider UDP Critical 2 class. This leaves the other non-priority service provider class (Critical 1) available for control plane applications, such as network control and call signaling, along with TCP-based transactional data applications. The figure shows the per-class re-marking requirements from the CE edge to gain access to the classes within the four-class service-provider model, with Cisco TelePresence assigned to the service-provider real-time class, along with voice.

In this example, Cisco TelePresence traffic must be re-marked on the CE egress edge to CS5 to gain access to the service provider real-time class. Also, broadcast video must be re-marked to CS2 to assign it to the service provider UDP class (Critical 2). Similarly, multimedia conferencing and multimedia streaming must be re-marked to AF2 to assign these also to the service-provider UDP class. Correspondingly, transactional data traffic must be re-marked to AF3 to gain access to the service-provider TCP class (Critical 1). All other traffic does not require re-marking to gain admission to the desired classes; this includes bulk and scavenger, because these default to the service provider best-effort class without any explicit re-marking.

Additionally, the relative per-class bandwidth allocations must be aligned, so that the enterprise CE edge queuing policies are consistent with the service-provider PE edge queuing policies to ensure compatible PHBs.



The six-class service-provider model is illustrated in the figure. In this model, there is a real-time class, a default best-effort class, a "less-than-best-effort" scavenger class, and three additional non-priority traffic classes. To illustrate more design options, Cisco TelePresence is assigned to a non-priority service provider class in this example; but of course, Cisco TelePresence can also be assigned (in combination with voice) to the service-provider real-time class.

In this case, the enterprise administrator can dedicate one of the non-priority classes (such as the service provider Critical 1) for Cisco TelePresence. It bears repeating that it is not recommended to assign Cisco TelePresence in conjunction with any unbounded application into a single service-provider class, because the other application could potentially cause the combined class to congest, resulting in Cisco TelePresence drops and loss of call quality.

This leaves two additional non-priority classes, which again allows the administrator to separate TCP-based applications from UDP-based applications. Specifically, broadcast video, multimedia conferencing, multimedia streaming, and OAM traffic can all be assigned to the service provider UDP class, Critical 3. This leaves the other non-priority service-provider class, Critical 2, available for control plane applications, such as network control and call signaling, along with TCP-based transactional data applications. The figure shows the per-class re-marking requirements from the CE edge to gain access to the classes within the six-class service provider model, with Cisco TelePresence assigned to a non-priority service-provider class.

As shown in the figure, in this second example Cisco TelePresence traffic does not need to be re-marked to gain access to the dedicated, non-priority service-provider class to which it is assigned (Critical 1). However, as before, Broadcast Video must be re-marked to CS2 to assign it to the service-provider UDP class, Critical 3; multimedia conferencing and multimedia streaming must be re-marked to AF2 to assign these also to the service-provider UDP class. Correspondingly, transactional data traffic must be re-marked to AF3 to gain access into the service-provider TCP class, Critical 2. All other traffic does not require re-marking to gain admission to the desired classes. However, it may be noted that bulk and scavenger no longer default to the service provider best-effort class, but now default to the service-provider scavenger class, which is the desired policy to bind these applications that potentially use much bandwidth. Additionally, the relative per-class bandwidth allocations again need to be aligned, so that the enterprise CE edge queuing policies are consistent with the PE edge queuing policies of the service provider.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- Before deploying QoS on your network, you should complete the following tasks:
  - Strategically define the business objectives to be achieved via QoS.
  - Analyze service-level requirements of the various traffic classes to be provisioned.
  - Design and test your QoS policies.
- By clearly defining organizational objectives and using the Cisco modified RFC 4594-based marking recommendations as guide, you can determine how many traffic classes you need and what those classes should be.
- Each enterprise traffic class has delay, jitter, packet loss, and bandwidth requirements.

Summary (Cont.)

- Because the Cisco QoS toolset provides many QoS design and deployment options, a few succinct design principles for classification, policing and markdown, and queuing and dropping can help simplify strategic QoS deployments. QoS classification best practice is achieved by:
  - Classifying and marking traffic as close to the source as possible.
  - Using DSCP markings whenever possible.
  - Following standards-based DSCP PHB markings to ensure interoperability and future expansion.
- Different enterprise traffic classes must be mapped into the specific traffic classes offered by the service provider.

Lesson 2

Deploying End-to-End QoS

Overview

When using public transportation, a traveler may benefit from contractual commitments from the transportation provider; for example, a guarantee from an airline that 95 percent of their flights will arrive within 5 minutes of the scheduled time. The commitments may include other parameters or metrics such as the number of stops en route. The more competitive the market for the particular service, the more comprehensive and tighter the commitments, or service level agreements (SLAs), that are offered.

In the same way, the increased competition between IP service providers and the heightened importance of IP to business operations has led to an increased demand and supply of IP services with tighter SLAs for IP performance.

The DiffServ architecture enables IP networks to be engineered to support tight SLA commitments. This lesson describes how the various quality of service (QoS) tools that we have discussed can be deployed in an end-to-end manner to achieve an end-to-end SLA.

Objective

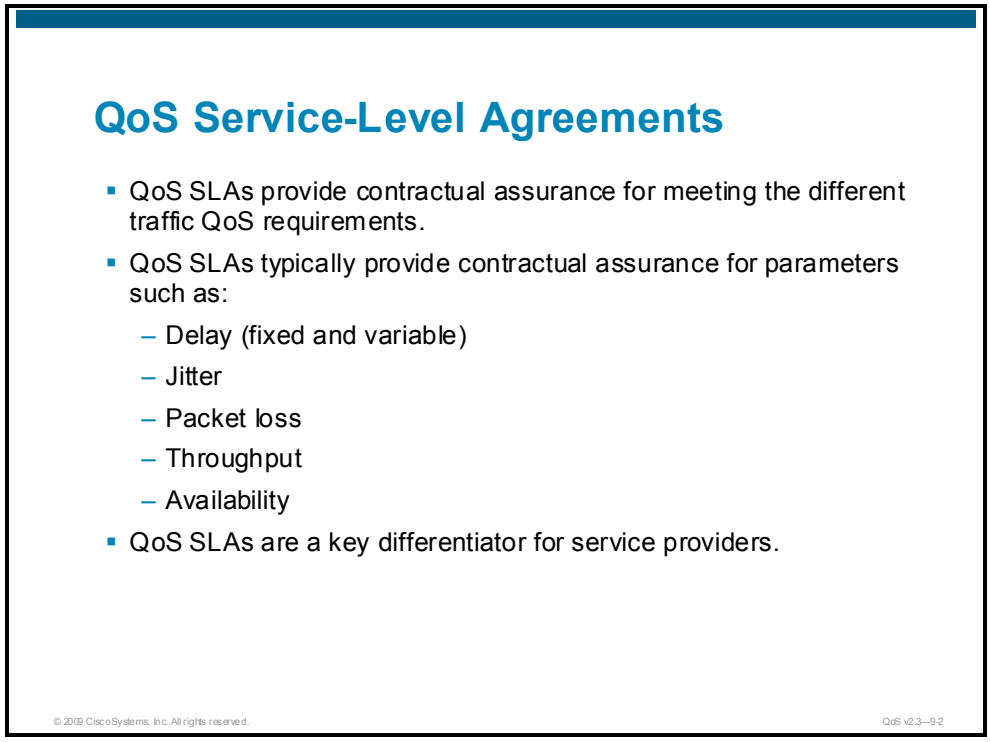
Upon completing this lesson, you will be able to describe the set of QoS mechanisms that are used to implement Cisco end-to-end QoS best practices in a typical enterprise network connected through a service provider providing Layer 3 IP services. This ability includes being able to meet these objectives:

- Describe QoS SLA
- Explain the typical network requirements within each functional block (campus LAN, WAN edge, service provider backbone, and branch) that makes up an end-to-end network
- Explain the best-practice QoS implementations and configurations within a campus LAN
- Describe access edge trust models
- Describe unique considerations for branch router QoS design
- Explain the best-practice QoS implementations and configurations on WAN CE and PE routers
- Explain the best-practice QoS implementations and configurations on the service provider IP core and PE routers

- Describe QoS design principles and designs to achieve end-to-end service levels over MPLS VPNs
- List general recommendations for end-to-end QoS

QoS Service Level Agreements

This topic describes QoS SLAs and provides some SLA examples.



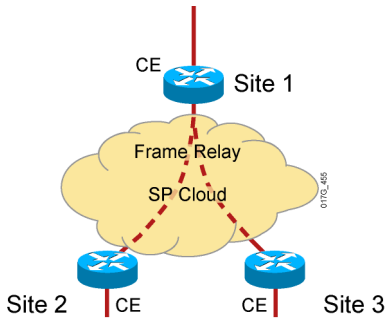
An SLA specifies the delivery and pricing of numerous service levels and spells out penalties for shortfalls. SLAs can cover an assortment of data services such as Frame Relay, leased lines, Internet access, web hosting, and so on. The best way to understand an SLA is to break it into two activities: negotiating the technology agreement and verifying compliance with the agreement.

To support integrated voice, video, and data services, service providers are under increasing pressure to offer differentiated service levels to their customers, often in conjunction with SLAs that provide contractual assurance for meeting the different traffic QoS requirements. A QoS SLA typically provides contractual assurance for parameters such as delay, jitter, packet loss, throughput, and availability.

With the rapid growth of new multimedia real-time applications such as IP telephony, web conferencing, and e-learning, QoS SLA is becoming a key service differentiator for service providers.

Enterprise Network with Traditional Layer 2 Service

- Provider sells the customer a Layer 2 service
- Point-to-point SLA from the provider
- Enterprise WAN likely to get congested
- QoS required for voice, video, data integration
- Service provider is not involved in QoS



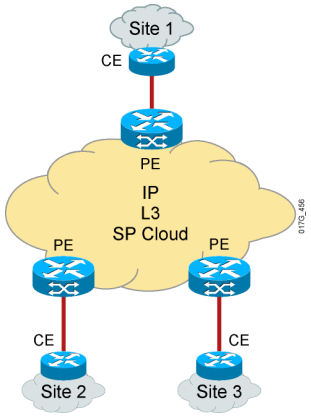
This figure illustrates a service provider providing only Layer 2 services to the enterprise customer. The customer edge (CE) routers at the various customer sites are interconnected by Frame Relay virtual circuits (VCs). These VCs can be fully meshed, partially meshed, or set up as hub-and-spokes, depending on the customer requirements.

In this environment, the service provider is responsible only for the end-to-end Layer 2 VC connections. The service provider provides only a point-to-point SLA guarantee for each VC connection, and is not involved with providing QoS to the customer.

To provide QoS for voice, video, and data integration over the Frame Relay VCs, the customer must configure the proper QoS mechanisms such as traffic shaping, low-latency queuing (LLQ), FRF.12, and compressed Real-Time Transport Protocol (cRTP) at the WAN customer edge (CE) routers, because the Frame Relay WAN link is likely to become congested.

Enterprise Network with IP Service

- Customer buys Layer 3 service from the provider
- Point-to-cloud SLA from provider for conforming traffic
- Enterprise WAN likely to get congested
- Service provider is involved in QoS

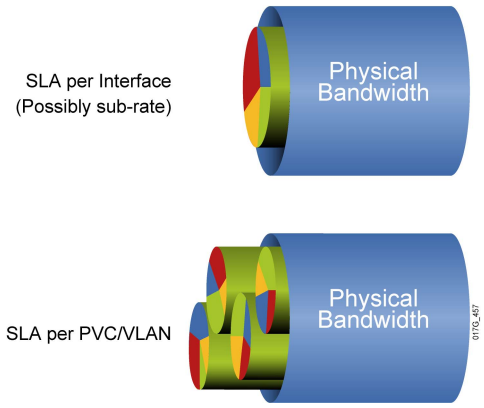


The figure illustrates a service provider that is providing Layer 3 services to the enterprise customer. The CE routers at the various customer sites connect to the provider edge (PE) of the service provider router. From a particular customer site perspective, every IP address that is not located on-site is reachable via the service provider IP backbone network.

In this environment, the service provider can provide value-added IP services to the customer by providing point-to-cloud SLAs for the conforming traffic from the customer. An SLA can, for example, divide customer traffic at the network edge into controlled latency, controlled load 1, and controlled load 2 classes, and then provide QoS assurances to each traffic class conforming to the contractual rate over a DiffServ IP backbone. For all nonconforming (exceeding) traffic, the service provider can re-mark and deliver all nonconforming traffic with best-effort service.

Know the SLA Offered by Your SP

- SLA typically includes three to five classes.
- Real-time traffic gets fixed bandwidth allocation.
- Data traffic gets variable bandwidth allocation with minimum guarantee.
- Additional classes not visible to customer may exist at the edge.

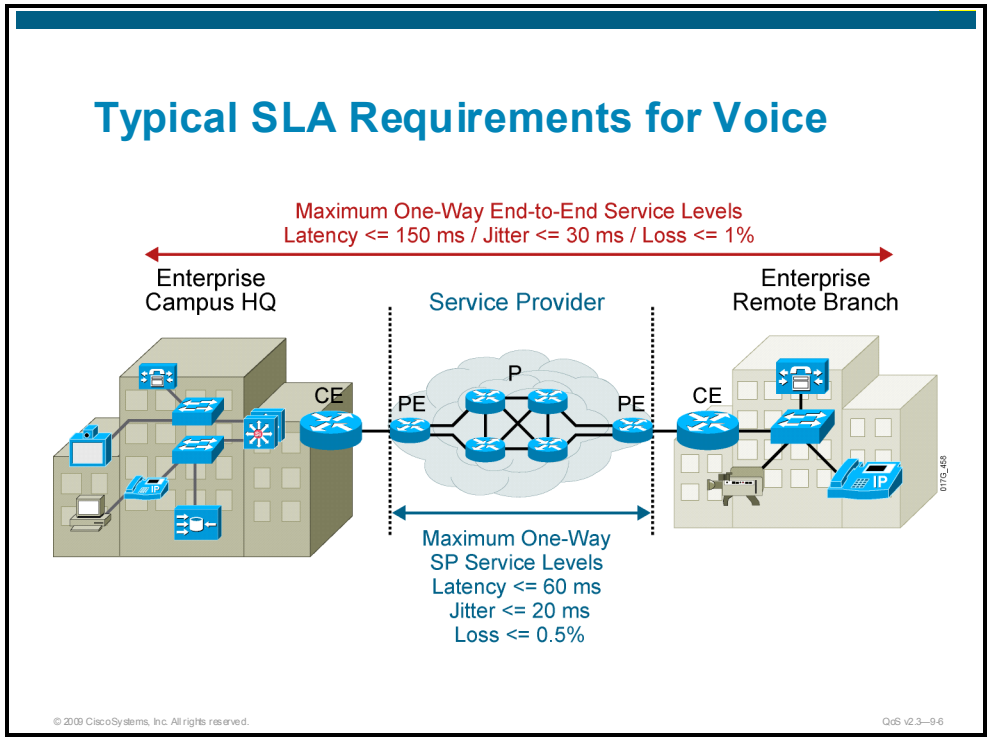


A typical QoS SLA offered by most service providers typically includes three to five traffic classes; for example, a real-time traffic class, a mission-critical data traffic class, one or two other data traffic classes, and a best-effort traffic class. The SLA for the real-time traffic class should be guaranteed a fixed maximum bandwidth, while the data traffic classes should be guaranteed a minimum bandwidth. Typically, the bandwidth allocation is configured as a percentage of the interface bandwidth. Each traffic class can also have a latency, delay, jitter, and packet-loss guarantee.

Between the CE and PE, there may be additional traffic classes that are used by the service providers only. For example, there may be a management traffic class for traffic such as Telnet or Simple Network Management Protocol (SNMP) from the service provider to the service provider-managed CE routers.

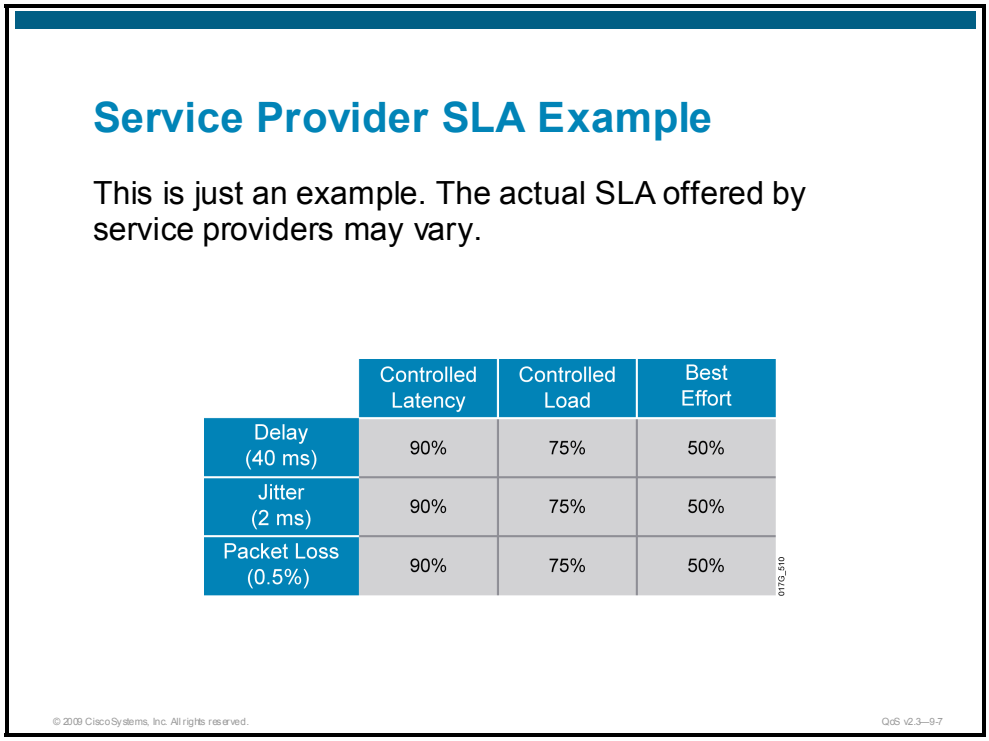
If a single physical interface is serving only one customer, the SLA is typically set up per interface. To provide easy bandwidth upgrades, service providers often install a high-speed link to the customer and then offer a sub-rate access.

If a single physical interface is serving many different customers, the SLA is typically set up per-permanent virtual circuit (PVC) or per-VLAN. To provide easy bandwidth upgrades, the service provider often installs a high-speed link to the customer and then offers sub-rate access.



To meet QoS requirements for different traffic types, both the enterprise and the service provider must implement the proper QoS mechanisms to provide end-to-end QoS for the packets traversing a service provider IP network. In the figure, the enterprise headquarters and the enterprise branch office are connected to a service provider that is providing Layer 3 services.

In this example, the service provider is providing an SLA for voice traffic with a latency of 60 ms or less, a jitter of 20 ms or less, and a packet loss of 0.5 percent or less. To meet the end-to-end QoS requirements for voice packets, the entire enterprise network must contribute less than 90 ms of delay—that is, 90 ms (enterprise network) + 60 ms (service provider network) <= 150 ms total one way delay. Similarly, jitter must be less than 10 ms—that is, 10 ms + 20 ms <= 30 ms total one way jitter. Finally, packet loss must be less than 0.5 percent—that is, 0.5 percent + 0.5 percent <= 1.0 percent total packet loss.



The figure shows an example of a typical QoS SLA from an IP service provider. In this example, the service provider offers three service classes to the customer: controlled latency, controlled load, and best effort. The SLA guarantees in this example include the following:

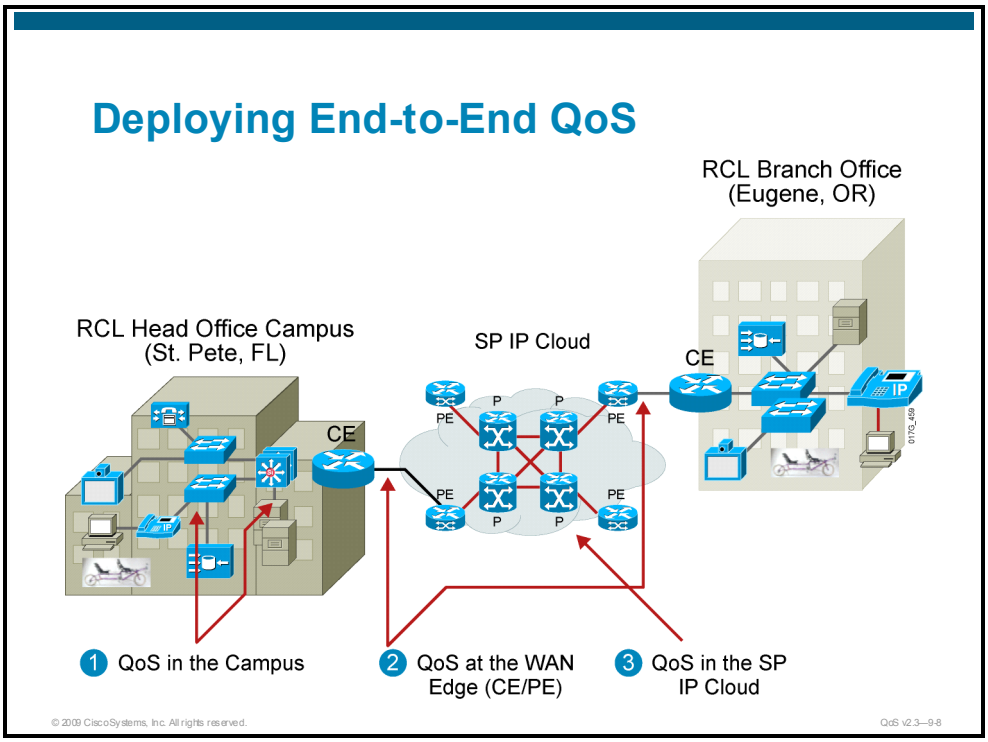
- For the controlled latency class:
  - A one-way delay of 40 ms that is guaranteed 90 percent of the time
  - A jitter of 2 ms that is guaranteed 90 percent of the time
  - A packet loss of 0.5 percent that is guaranteed 90 percent of the time
- For the controlled load class:
  - A one-way delay of 40 ms that is guaranteed 75 percent of the time
  - A jitter of 2 ms that is guaranteed 75 percent of the time
  - A packet loss of 0.5 percent that is guaranteed 75 percent of the time
- For the best-effort class:
  - A one-way delay of 40 ms that is guaranteed 50 percent of the time
  - A jitter of 2 ms that is guaranteed 50 percent of the time
  - A packet loss of 0.5 percent that is guaranteed 50 percent of the time

This is just an example. Actual SLA offered by service providers may vary.

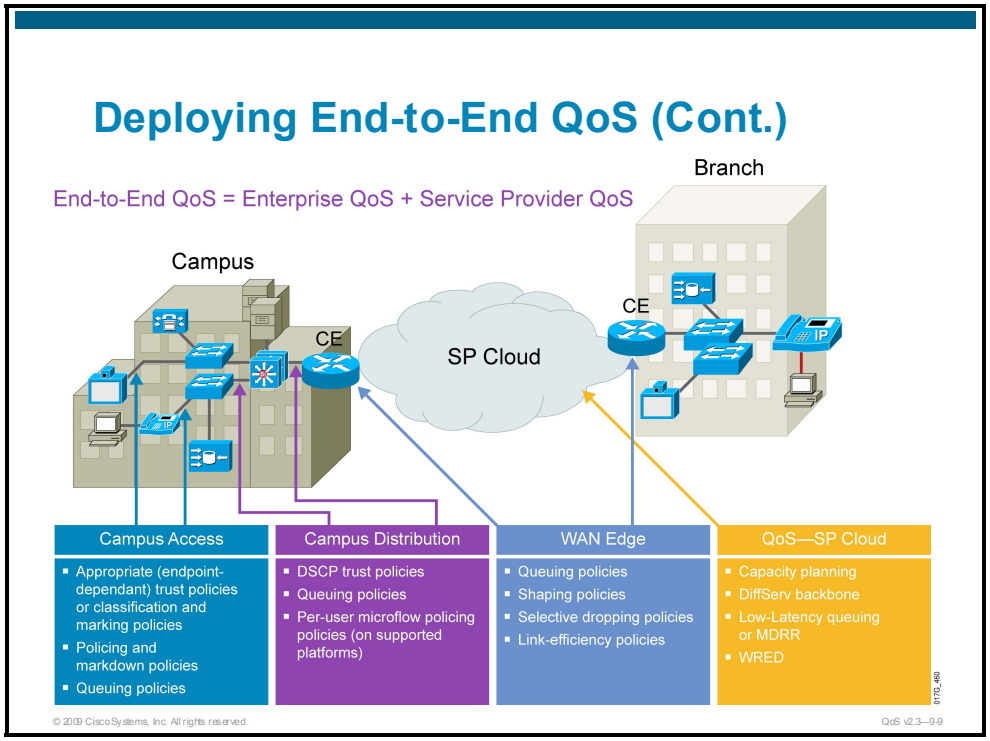
In the U.S. tier 1 Internet backbone, a typical round-trip time (RTT) delay between two service provider points of presence (POPs) is about 40 ms. It is typical for a service provider to offer a monthly average loss on its network of less than 1 percent.

Deploying End-to-End QoS

This topic describes the typical network requirements within each functional block which makes up an end-to-end network (headquarters, campus LAN, WAN edge, service provider backbone, and branch office).



To meet the QoS requirements for different traffic types, both the enterprise and the service provider must implement the proper QoS mechanisms to provide end-to-end QoS for the packets traversing a service provider network.

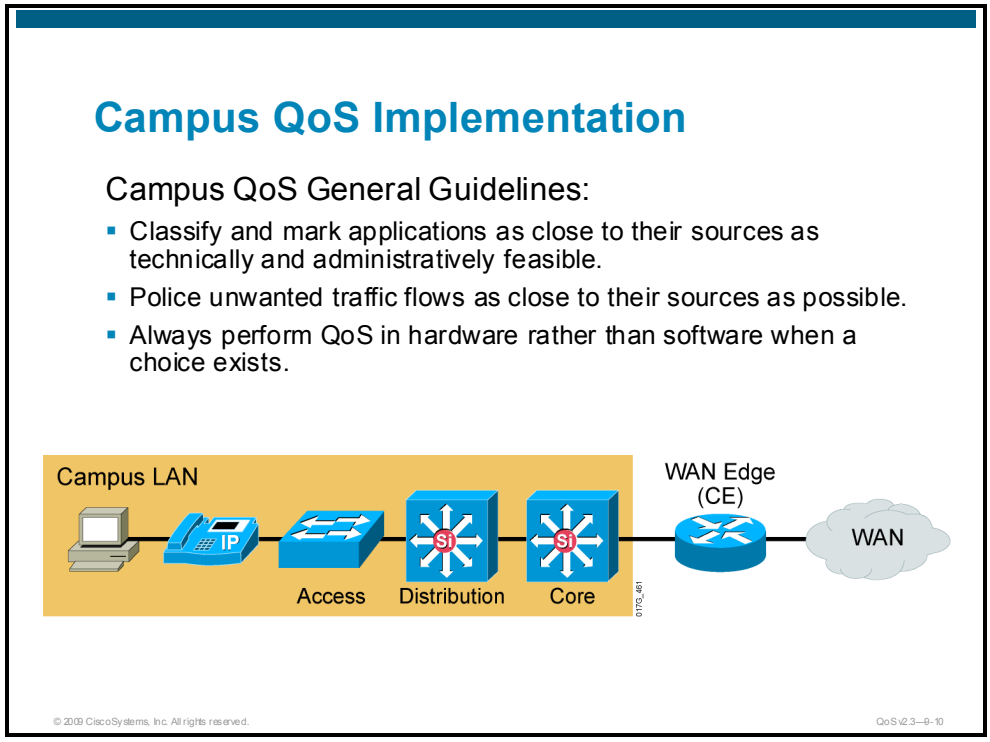


To provide end-to-end QoS, both the enterprise and the service provider must implement the proper QoS mechanisms to ensure the proper per-hop behavior (PHB) for each traffic class across the whole network. Until recently, QoS was not an issue in an enterprise campus network where bandwidth is plentiful. But as more applications such as IP telephony, interactive video (or IP videoconferencing, known also as IP/VC), e-learning, and mission-critical data applications are being implemented in the campus, it has become evident that buffer management, not just bandwidth, is an issue that must be addressed. QoS functions such as classification, scheduling, and provisioning are also now required within the campus to manage bandwidth and buffers to minimize loss, delay, and jitter.

The figure lists some of the requirements within the different building blocks that make up the end-to-end network. Most of the more complex QoS configurations occur at the WAN edge. In the IP core (service provider cloud), only queuing, such as LLQ with class-based weighted fair queuing (CBWFQ) or modified deficit round robin (MDRR), and weighted random early detection (WRED) should be required.

Enterprise Campus QoS General Guidelines

This topic describes some of the best-practice QoS implementations and configurations within the campus LAN.



Although network administrators sometimes equate QoS only with queuing, the QoS toolset extends considerably beyond queuing tools. Classification, marking, and policing are all important QoS functions that are optimally performed within the campus network, particularly at the access layer ingress edge (the access edge).

Three QoS design principles are important when deploying campus QoS policies:

- Classify and mark applications as close to their sources as technically and administratively feasible. This principle promotes end-to-end Differentiated Services PHBs. Sometimes endpoints can be trusted to set Class of Service (CoS) or Differentiated Services Code Point (DSCP) markings correctly, but this is not recommended because users can easily abuse provisioned QoS policies if permitted to mark their own traffic. For example, if DSCP Expedited Forwarding (EF) received priority services throughout the enterprise, users could easily configure the NIC on a PC to mark all traffic to DSCP EF, thus hijacking network priority queues to service their non-real time traffic. Such abuse could easily ruin the service quality of real-time applications, such as VoIP, throughout the enterprise. For this reason, the phrase "as close as... administratively feasible" is included in the design principle.
- Police unwanted traffic flows as close to their sources as possible. There is little sense in forwarding unwanted traffic only to police and drop it at a subsequent node. This is especially the case when the unwanted traffic is the result of Denial of Service (DoS) or worm attacks. Such attacks can cause network outages by overwhelming network device processors with traffic.

- Always perform QoS in hardware rather than software when a choice exists. Cisco IOS routers perform QoS in software. This places additional demands on the CPU, depending on the complexity and functionality of the policy. Cisco Catalyst switches, on the other hand, perform QoS in dedicated hardware ASICS and therefore do not tax their main CPUs to administer QoS policies. You can therefore apply complex QoS policies at gigabit or 10 Gigabit Ethernet line speeds in these switches.

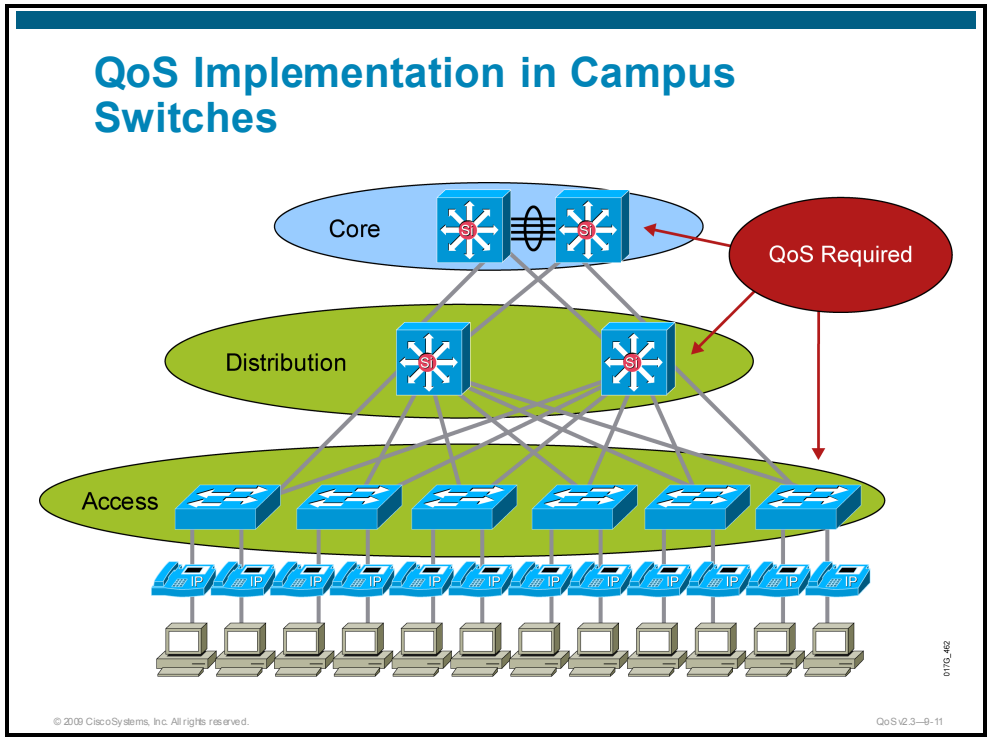
For these reasons, you should enable QoS policies such as classification and marking policies to establish and enforce trust boundaries as well as policers to protect against undesired flows at the access edge of the LAN.

Most campus links are underutilized. Some studies have shown that 95 percent of campus access layer links are utilized at less than 5 percent of their capacity. This means that you can design campus networks to accommodate oversubscription between access, distribution, and core layers. Oversubscription allows for uplinks to be utilized more efficiently, and more importantly, it reduces the overall cost of building the campus network. Common campus oversubscription values are 20:1 for the access-to-distribution layers and 4:1 for the distribution-to-core layers.

It is quite rare under normal operating conditions for campus networks to suffer congestion. If congestion does occur, it is usually momentary and not sustained, as at a WAN edge. However, critical applications like VoIP still require service guarantees regardless of network conditions.

The only way to provide service guarantees is to enable queuing at any node that has the potential for congestion, regardless of how rarely this may actually occur. The potential for congestion exists in campus uplinks because of oversubscription ratios and speed mismatches in campus downlinks (for example, Gigabit Ethernet to Fast Ethernet links). The only way to provision service guarantees in these cases is to enable queuing at these points.

Queuing helps to meet network requirements under normal operating conditions, but enabling QoS within the campus is even more critical under abnormal network conditions such as DoS or worm attacks. During such conditions, network traffic may increase exponentially until links are fully utilized. Without QoS, the worm-generated traffic drowns out applications and causes denial of service through unavailability. Enabling QoS policies within the campus maintains network availability by protecting and servicing critical applications such as VoIP and even best-effort traffic.



Access switches require the following QoS policies:

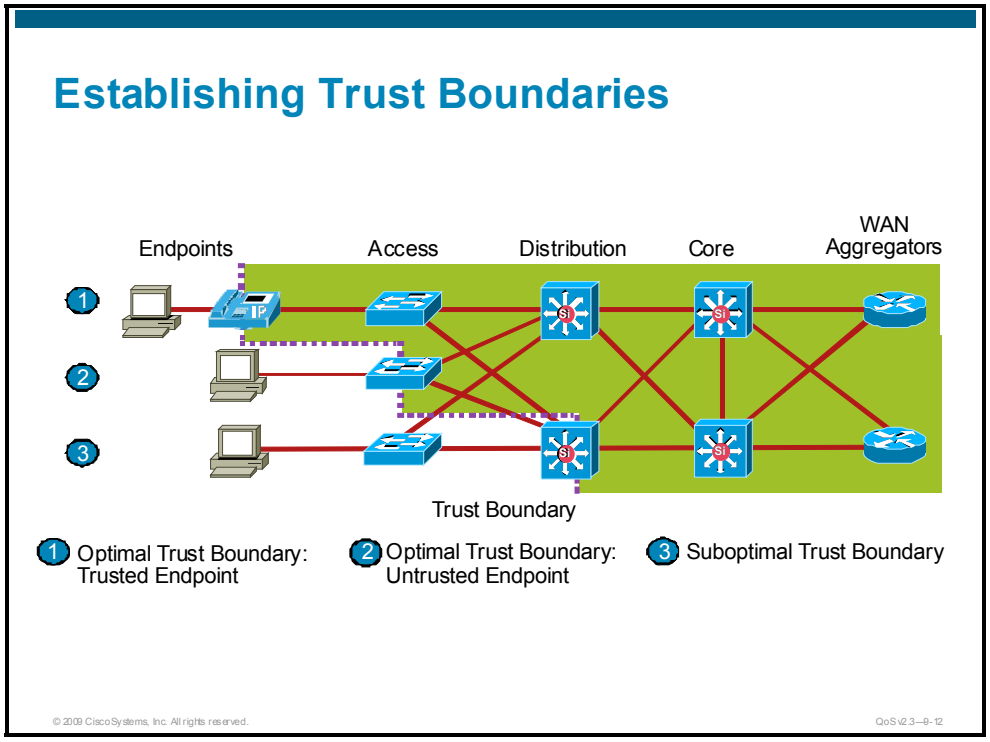
- Appropriate (endpoint-dependant) trust policies or classification and marking policies
- Policing and markdown policies
- Queuing policies

Distribution and core switches require the following QoS policies:

- DSCP trust policies
- Queuing policies
- Optional per-user microflow policing policies (only on supported platforms)

Access Edge Trust Models

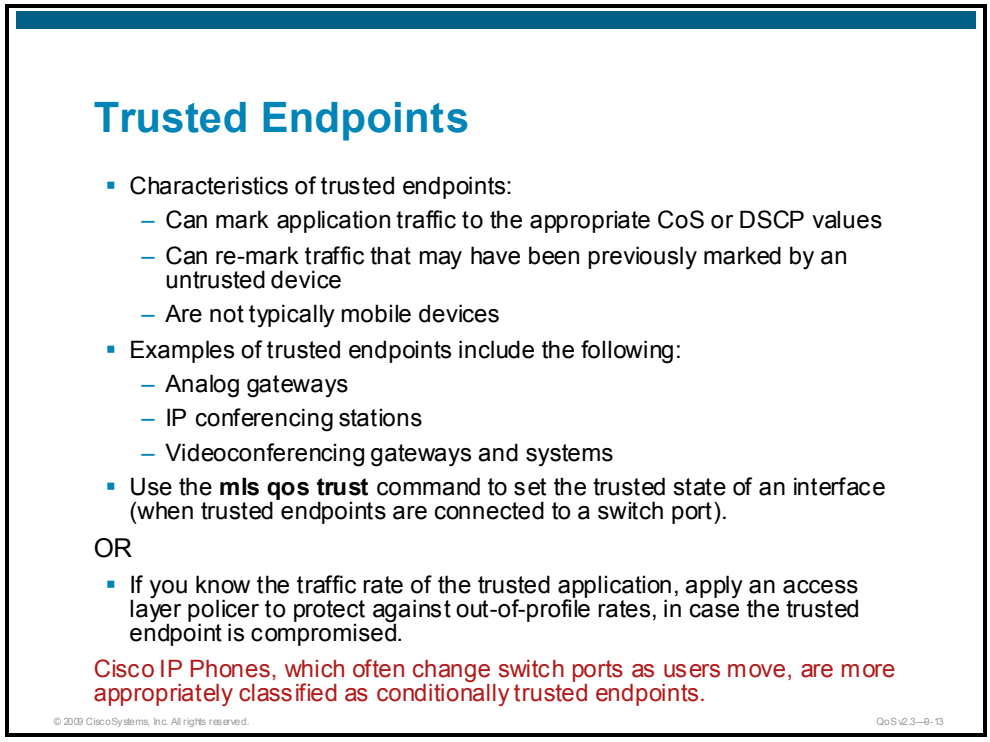
This topic describes access edge trust models, including trusted endpoints, untrusted endpoints, and conditionally trusted endpoints.



The primary function of access edge policies is to establish and enforce trust boundaries. A trust boundary is the point within the network where markings such as CoS or DSCP begin to be accepted. Previously set markings are overridden as required at the trust boundary. You should enforce trust boundaries as close to the endpoints as technically and administratively possible, as shown in the figure.

The definition of the trust boundary depends on the capabilities of the endpoints that are being connected to the access edge of the LAN. The following are the three main categories of endpoints as they relate to trust boundaries:

- Trusted endpoints
- Untrusted endpoints
- Conditionally trusted endpoints



Trusted endpoints have the capabilities and intelligence to mark application traffic to the appropriate CoS or DSCP values. Trusted endpoints also have the ability to re-mark traffic that may have been previously marked by an untrusted device. Trusted endpoints are not typically mobile devices, which means that the switch port into which they are plugged does not usually change.

Note	Cisco IP Phones, which often change switch ports as users move, are more appropriately classified as conditionally trusted endpoints.

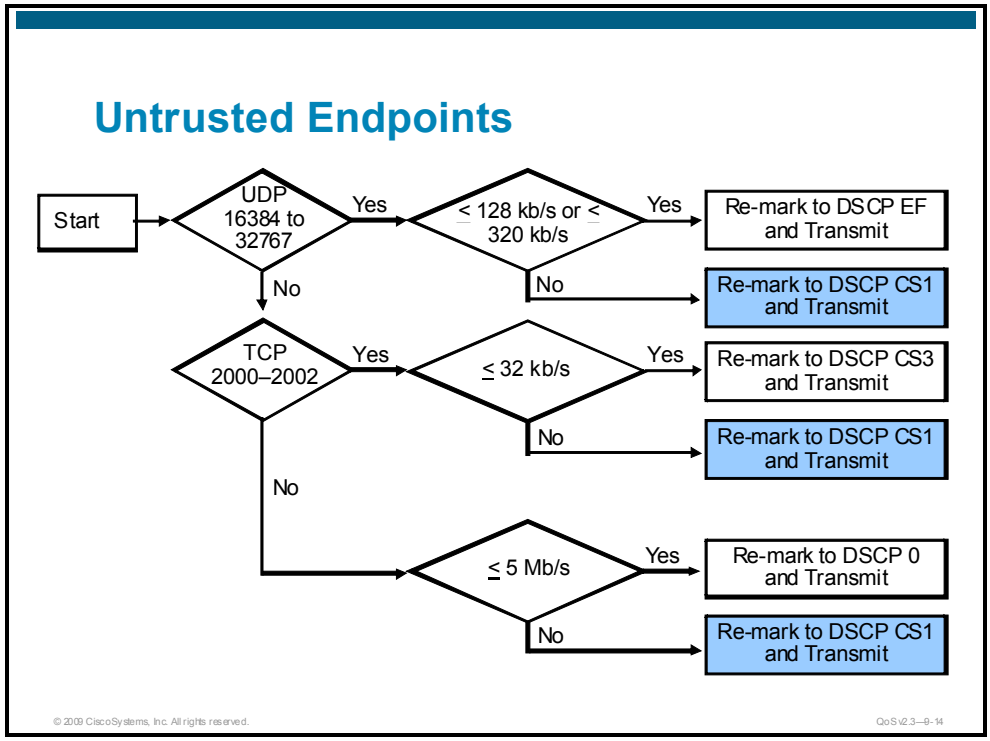
Examples of trusted endpoints include the following:

- **Analog gateways:** These devices connect analog devices such as fax machines, modems, TDD TTYs, and analog phones to the VoIP network, such that the analog signals can be packetized and transmitted over the IP network. Cisco VG224 Voice Gateway and Cisco VG248 Analog Phone Gateway are examples of analog gateways.
- **IP conferencing stations:** These devices are specialized IP Phones with 360-degree microphones and advanced speakerphones designed for meeting room VoIP conferencing. Examples of such devices include the Cisco Unified IP Conference Station 7935 and 7936.
- **IP/VC gateways and systems:** These devices transmit interactive video across the IP network. Examples of such devices capable of setting DSCP markings include the Cisco Unified Videoconferencing systems 3511, 3521, 3526 and 3540. If a video-conferencing device does not have the ability to set DSCP markings correctly, it should be treated as an untrusted device.
- **Video surveillance units**: These third-party devices are used for security and remote monitoring purposes over an IP (as opposed to a closed-circuit) network. They may support DSCP marking, in which case they may be considered trusted endpoints.

- **Servers:** Certain servers, within the data center or otherwise, might be capable of correctly marking their traffic on their NICs. In such cases, you can choose to trust such markings. However, enforcing such a trust boundary requires cooperation between network administrators and system or server administrators, an alliance that is often fragile. Additionally, the majority of DoS and worm attacks target servers. Infected servers not only might send profuse amounts of traffic onto the network, but, in such cases, they might do so with trusted markings. There is no hard-and-fast rule that applies to every situation. Some administrators prefer to trust certain servers (such as Cisco Unified Communications Manager) to provide services, due to the large number of ports that may be in use, rather than administer complex access control lists (ACLs).
- **Wireless access points:** Some wireless access points (APs) have the ability to mark or remark 802.1p CoS and DSCP values and therefore qualify as trusted endpoints. Examples include Cisco Aironet 350, 1100 and 1200 series APs.
- **Wireless IP Phones:** Mobile wireless IP Phones can mark DSCP values for VoIP and call signaling and pass these on to the wireless AP with which they are associated. The Cisco Unified Wireless IP phone 7920G is an example.

When trusted endpoints are connected to a switch port, all that is typically required is using the **mls qos trust dscp** command in interface configuration mode.

Optionally, if you know the traffic rate of the trusted application, you can apply an access layer policer to protect against out-of-profile rates, in case the trusted endpoint is somehow compromised. For example, consider the case of an IP videoconferencing (IP/VC) station that transmits 384 kb/s of video (not including Layer 2-4 overhead) and correctly marks this traffic to DSCP AF41. An access edge ingress policer could be applied to the switch port to which this IP/VC station is connected and be configured to trust up to 500 kb/s, allowing for Layer 2-4 overhead and policer granularity of interactive video traffic marked AF41. Excess traffic could be marked to CS1. Such a policy prevents network abuse if another device is inserted, perhaps via a hub, into the path, or if the trusted endpoint itself becomes compromised.



As previously mentioned, trusting end-user PCs is generally a bad idea because newer operating systems, such as Windows XP and Linux, make it relatively easy to set CoS or DSCP markings on PC NICs. Such markings may be set deliberately or even inadvertently. In either case, improperly set QoS markings can affect the service levels of multiple users within the enterprise and make troubleshooting very difficult. Also, marking application traffic on server NICs has disadvantages (as discussed in the previous section) that may make it preferable to treat PCs as untrusted devices. While client PCs and data center servers are related and complimentary, they also have unique considerations that affect their classification and marking policies. Some of these considerations follow:

Untrusted PC + SoftPhone with Scavenger-Class QoS

Cisco generally recommends not trusting end-user PC traffic. However, some PCs may be running applications that critically require QoS treatment. A classic example is a PC running the Cisco IP SoftPhone application. In such a case, the critical application must be identified using ACLs and marked or re-marked at the access edge. Re-marking can be done with either an MLS QoS **set ip dscp** command or with a policer.

A policer is recommended in this case, because limits on the amount of traffic being marked can then be imposed to prevent abuse. A Cisco SoftPhone can use regular G.711 codecs, in which case 128 kb/s is adequate, or they can be configured use a G.722 (wide codec), in which case 320 kb/s is required. The tighter the policer, the better, provided that adequate bandwidth has been allocated for application requirements.

Additionally, you can explicitly define the UDP ports used by Cisco SoftPhone within the application as opposed to simply choosing random ports within the UDP range of 16383 to 32767. This is recommended because it allows for a more granular ACL to match legitimate Cisco SoftPhone traffic, thereby tightening the overall security of the policy.

Note	In this context, “SoftPhone” can refer to any PC-based IP telephony application, including Cisco IP Communicator and similar products.

The logic of such an access edge policer marking Cisco SoftPhone traffic from an untrusted PC endpoint is shown in the figure. The syntax for implementing such a policer may vary slightly from platform to platform.

Untrusted Server with Scavenger-Class QoS

Servers as well as PCs are subject to attack and infection by worms and viruses, so they should also be policed as to the amounts of traffic they admit onto the network. The values are greater than PC endpoints, so you should profile traffic patterns from servers to establish a baseline of normal and abnormal behavior.

For example, assume that a single server is running multiple applications, in this case, SAP (TCP ports 3200-3203 and 3600), Lotus Notes (TCP port 1352), and Internet Message Access Protocol (IMAP) (TCP ports 143 and 220). SAP is considered a mission-critical application, and until call signaling marking on IP telephony equipment fully migrates from DSCP AF31 to CS3, it should be marked to DSCP 25. Lotus Notes is classed as a transactional data application and should be marked to DSCP AF21. IMAP is considered a bulk application and should be marked to DSCP AF11.

Application baselining has shown that 95 percent of the traffic rates for SAP, Lotus Notes and IMAP are less than 15 Mb/s, 35 Mb/s and 50 Mb/s, respectively. To ensure that no other traffic comes from the server, a final policer to catch any other type traffic is included. In the event of legitimate traffic that temporarily exceeds these values, no dropping or re-ordering of packets occurs. However, should this server become infected and begin sending sustained traffic in excess of these normal rates, the excess is subject to aggressive dropping in the event of link congestion.

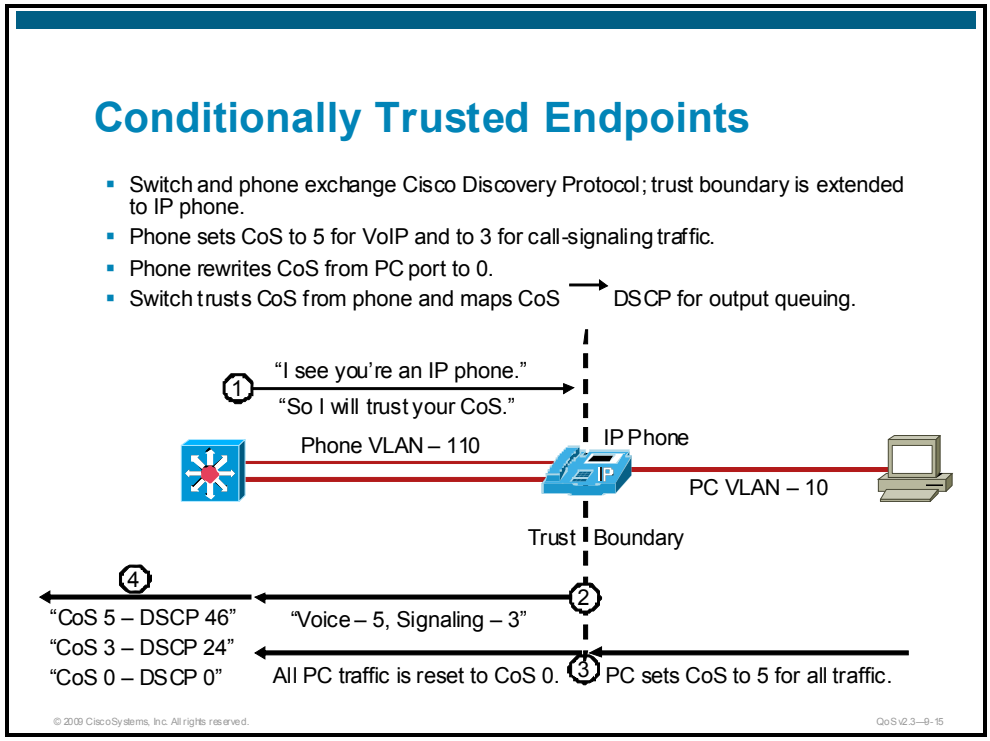
Remember that when deploying QoS designs for untrusted servers, the applications are usually identified by source ports, rather than destination ports (as is the case with client-to-server ACLs). Thus, the ACL becomes the following:

```
permit [tcp | udp] any [eq | range] any
```

as opposed to:

```
permit [tcp | udp] any any [eq | range]
```

This is a subtle but critical difference.



One of the main business advantages of IP telephony is the simplicity and related cost savings of user additions, moves, and changes. To move, users simply pick up their IP phones, plug them in at their new locations, and carry on business as usual. If their infrastructure supports inline power, it is literally a matter of unplugging a single RJ-45 cable and plugging it in at the new location.

IP phones are trusted devices, while PCs are not. This can be a problem when provisioning trust in a mobile environment. Consider the following example: Port A is configured to trust the endpoint connected to it, which initially is an IP phone. Port B is configured not to trust the endpoint connected to it, which initially is a PC. Because of a move, these endpoints are plugged into the opposite ports. This breaks the VoIP quality of calls made from the IP phone (now plugged into untrusted Port B) and opens the network up for unintentional or deliberate abuse of provisioned QoS by the PC (now plugged into the trusted Port A).

One solution is to place a call to the networking help desk when the move is scheduled, so that the switch ports can be reconfigured to trust or untrust the endpoints, as required. However, this approach dampens the mobility business advantage of IP telephony, because manual network administration is then required to complete the move.

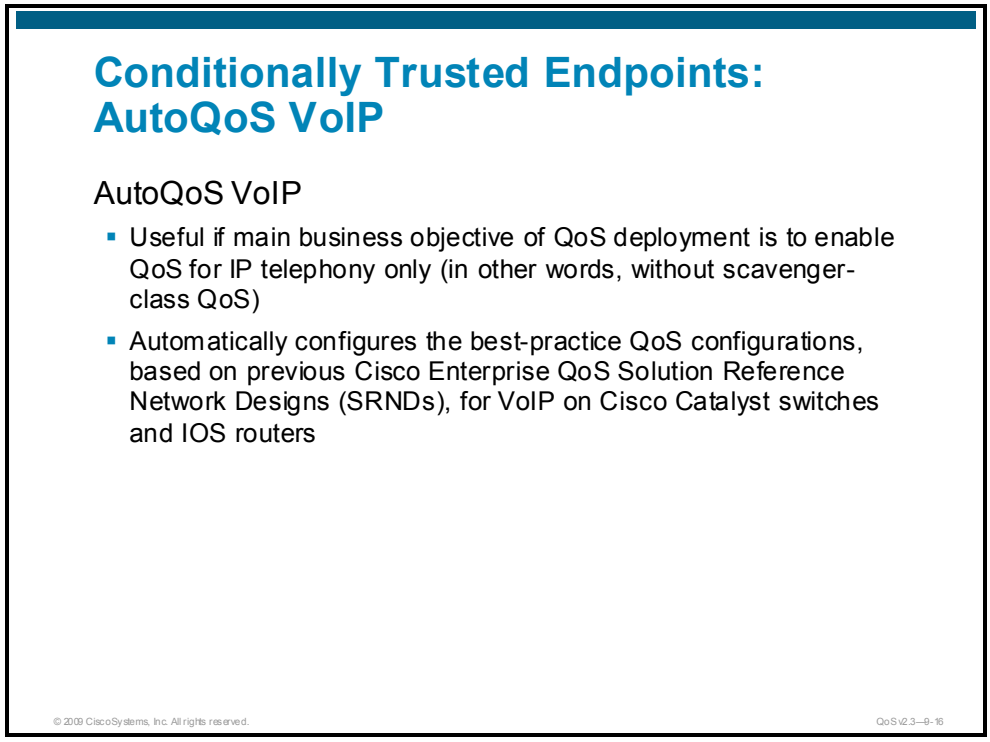
Another solution is to have an intelligent exchange of information between the switch and the devices plugged into its ports. If the switch discovers a device that is trustworthy, it can extend trust to the device dynamically.

Cisco IP phones use the latter solution. In the current Cisco implementation, the intelligent exchange of information is performed using Cisco Discovery Protocol, a lightweight, proprietary protocol engineered to perform neighbor discovery. It was never intended as a security or authentication protocol. Therefore, to improve the security of conditional trust extension, the next generation of Cisco IP telephony products will incorporate the use of advanced protocols to perform authentication.

The figure shows a conditional trust boundary extension granted to an IP Phone that has passed a Cisco Discovery Protocol exchange. The sequence shown is as follows:

1. The switch and the phone exchange Cisco Discovery Protocol; the trust boundary is extended to the IP phone.
2. The phone sets CoS to 5 for VoIP and to 3 for call-signaling traffic.
3. The phone rewrites CoS from the PC to 0.
4. The switch trusts CoS from the phone and maps CoS to DSCP for output queuing.

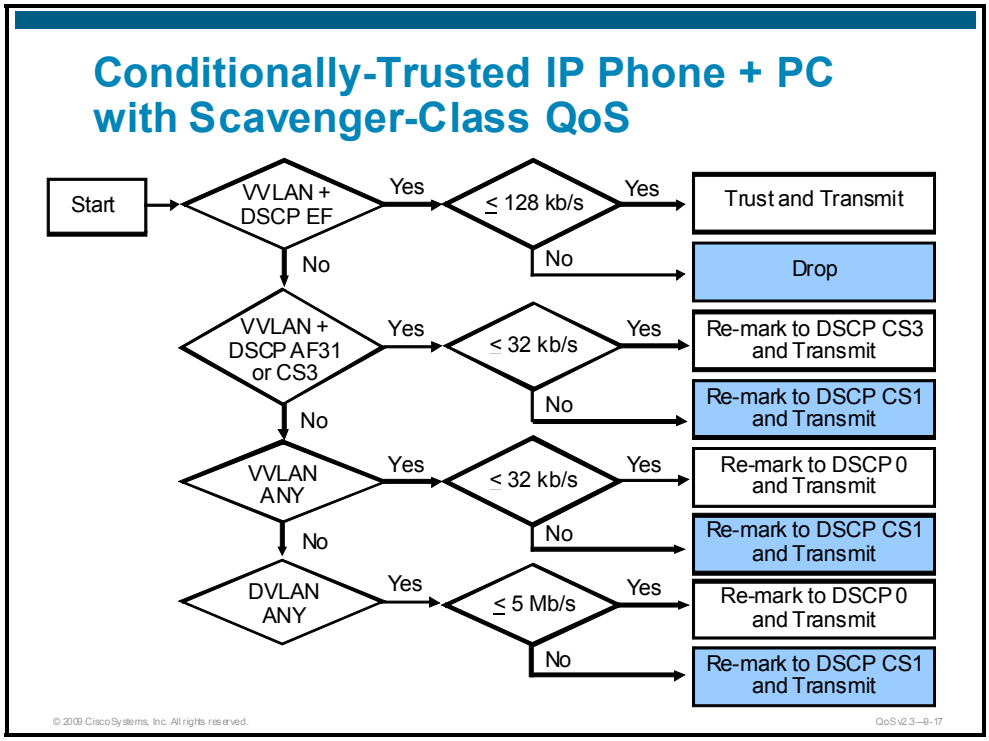
Note For an overview of some of the main Cisco IP phones and their impact on access edge QoS design, refer to the “Campus QoS Design” section of the *Enterprise QoS Solution Reference Network Design Guide* at the following URL: http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND/QoSDesign.html#wp998284



If the main business objective of the QoS deployment is to enable QoS for IP telephony only (that is, without scavenger-class QoS), you may choose to take advantage of the Cisco AutoQoS VoIP feature. Cisco AutoQoS VoIP is essentially an intelligent macro that enables you to enter one or two simple Cisco AutoQoS commands to enable all the appropriate features for the recommended QoS settings for a VoIP and IP telephony for a specific platform and a specific interface.

Cisco AutoQoS VoIP automatically configures the best-practice QoS configurations, based on previous Cisco Enterprise QoS Solution Reference Network Designs (SRNDs), for VoIP on Cisco Catalyst switches and Cisco IOS routers. For example, on Cisco Catalyst switches, AutoQoS performs the following automatically:

- Enforces a conditional-trust boundary with any attached Cisco IP phones
- Enforces a trust boundary on Catalyst switch access ports and uplinks or downlinks
- Modifies CoS-to-DSCP (and IP Precedence-to-DSCP) mappings, as required
- Enables Catalyst strict priority queuing for voice (CoS 5/DSCP EF) and preferential queuing for Call-Signaling traffic (CoS 3/DSCP CS3)
- Enables best-effort queuing for all other data (CoS 0/DSCP 0) traffic
- Modifies queue admission criteria (such as CoS-to-queue mapping)
- Modifies queue sizes and queue weights where required



Conditionally Trusted IP Phone Plus PC with Scavenger-Class QoS (Basic) Model

In the conditionally trusted IP phone plus PC with scavenger-class QoS basic model, trust of CoS markings is extended to Cisco Discovery Protocol verified IP phones. An additional layer of protection can be offered by access edge policers. As stated previously, the tighter the policers, the better, provided that adequate bandwidth is permitted for legitimate applications. The most granular policing can be achieved by the use of per-port per-VLAN policers.

Note	Currently, only the Catalyst 3550 Series family supports per-port per-VLAN policing as a feature. For platforms that do not yet support this feature, equivalent logic can be achieved by including subnet information within the access lists being referenced by the class maps.

For example, the peak amounts of legitimate traffic originating from the voice VLAN (VVLAN) on a per-port basis are as follows:

- 128 kb/s for voice traffic, marked CoS 5/DSCP EF (320 kb/s in the case of G.722 codecs)
- 32 kb/s for call signaling traffic (marked CoS 3/DSCP AF31 or CS3)
- 32 kb/s of best-effort services traffic (marked CoS 0)

There should not be any other traffic originating from the VVLAN; therefore, the policer can be configured to remark anything else from the VVLAN, because such traffic is considered illegitimate and indicative of an attack.

These policers can then be combined with a policer to meter traffic from the data VLAN (DVLAN), marking down traffic in excess of 5 percent (5 Mb/s for FE ports) to scavenger (CS1). The logic of these policers is shown in the figure.

Conditionally Trusted IP Phone Plus PC with Scavenger-Class QoS (Advanced) Model

Building on the previous model, you can add additional marking and policing for PC-based video-conferencing and multiple levels of data applications. Desktop videoconferencing applications use the same UDP port range by default as does Cisco IP SoftPhone. If the UDP ports used by the desktop videoconferencing application can be explicitly defined within the application, as with Cisco IP SoftPhone, you can use two policers: one for IP/VC and another for Cisco IP SoftPhone. Otherwise, a single policer covering the UDP port range of 16384 to 32767 is required, which would be provisioned for the worst-case scenario of legitimate traffic. In this case, this is the videoconferencing application requirement of 500 kb/s (for a 384 kb/s desktop IP/VC application), as compared to the Cisco IP SoftPhone requirement of 128 kb/s (or 320 kb/s for G.722 codecs).

Policer thresholds should be set according to the requirements of the video application. Some interactive video applications may have higher bandwidth requirements for their codecs.

You can add data VLAN policers to meter transactional data and bulk (High-Throughput) data flows. Each of these classes can be policed on ingress to the switch port to an in-profile amount, such as 5 percent each.

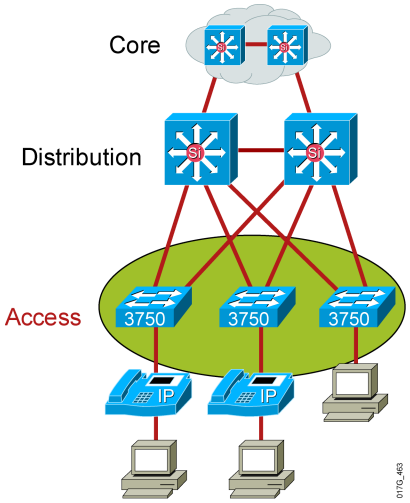
Note Because transactional data applications are interactive foreground applications requiring user input, it is highly unlikely that these types of applications will simultaneously generate 5 Mb/s each from a client PC. However, in the rare case that they do, these flows will be policed further by any per-user microflow policing policies that may be deployed on distribution layer Catalyst 6500 Series Supervisor 720s (PFC3s).

Another factor to remember is that certain Catalyst platforms allow only up to 8 policers per Fast Ethernet port. Therefore, the model presented here is made to conform to this constraint to make it more generic and modular. For this reason, a separate policer has not been defined for call signaling traffic from SoftPhone. An access list to identify such traffic could be included within the mission-critical data ACLs.

Note For more information on Access Edge Trust Models, refer to the “Campus QoS Design” section of the *Enterprise QoS Solution Reference Network Design Guide* on Cisco.com.

Catalyst 3750 Series Switch

- 1P3Q3T (recommended) or 4Q3T
- Priority queue (Queue 1)
  - Should be enabled in converged campus environment
- 3 WTD thresholds per queue
  - 2 configurable.
  - 1 nonconfigurable (set to queue-full state, 100%).
  - Queue 2: Set first threshold to 70% and second to 80%.
  - Queue 4: Set first threshold to 40% and accept default value for second and third thresholds (100%).
- QoS disabled globally by default
  - When QoS is enabled, all DSCP and CoS values are set by default to 0.
- SRR (shaped or shared)



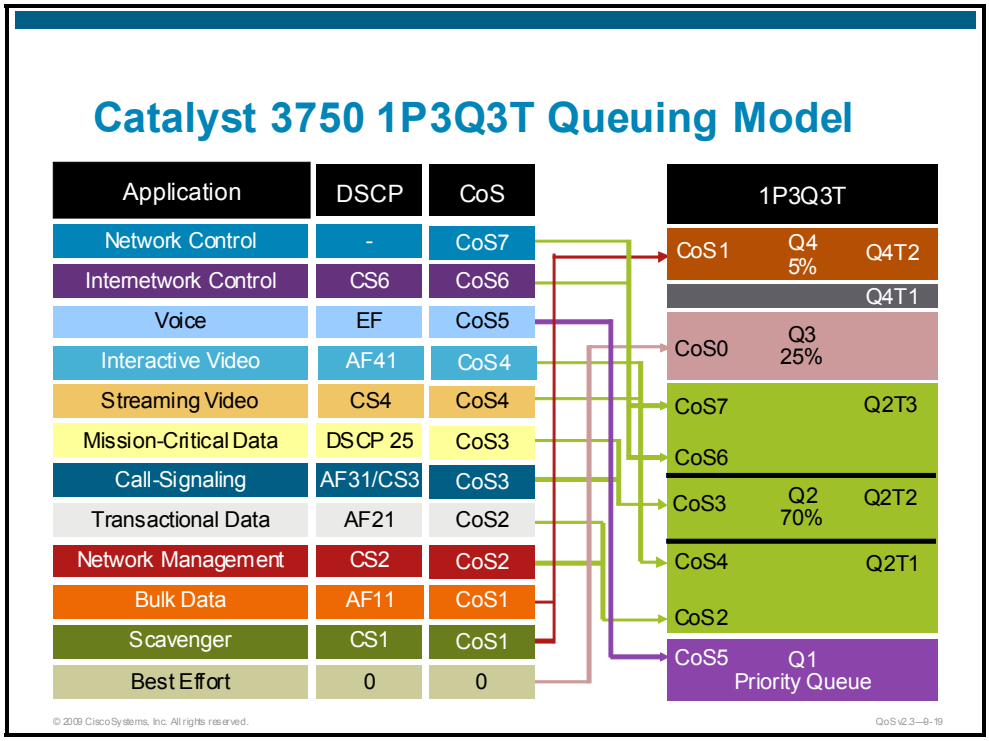
The Catalyst 3750 switch supports Layer 3 routing and may be found in either the access layer or the distribution layer. QoS is globally disabled by default on the Catalyst 3750 switch. While QoS is disabled, all frames and packets are passed through the switch unaltered (which is equivalent to a trust CoS and trust DSCP state on all ports). When QoS is globally enabled, however, all DSCP and CoS values are set by default to 0 (which is equivalent to an untrusted state on all ports). QoS must be enabled globally for configured policies to become effective.

The Catalyst 3750 switch supports four egress queues, which can be configured on a per-interface basis to operate in either 4Q3T or 1P3Q3T modes. Additionally, the Catalyst 3750 switch supports two queue-sets, allowing certain interfaces to be configured in one manner and others to be configured in a different manner. For example, some interfaces may be assigned to queue Set (qset) 1 operating in 4Q3T mode, while others may be assigned to queue Set 2 operating in 1P3Q3T mode.

The Catalyst 3750 switch has queue 1 as the optional priority queue. In a converged campus environment, it is recommended to enable the priority queue. The three remaining egress queues on the Catalyst 3750 switch are scheduled by a shaped round-robin (SRR) algorithm, which can be configured to operate in shaped mode or in shared mode.

Note The Catalyst 3750 switch also supports two configurable ingress queues (normal and expedite). Ingress scheduling, however, is rarely (if ever) required; it becomes enabled only if the combined input rates from any or all switch ports exceed the capacity of the switch fabric. Such cases are extremely difficult to achieve, even in controlled lab environments. In the extreme case where such a scenario develops in a production environment, the default settings of the ingress queues are acceptable to maintain VoIP quality and network availability.

The Catalyst 3750 switch supports three weighted tail drop (WTD) thresholds per queue. Two of these thresholds are configurable (explicit); the third is non-configurable (implicit), because it is set to the queue-full state (100 percent). The only queues that need to have these thresholds defined (away from the defaults) are queues 2 and 4. In queue 2, it is recommended to set the first threshold to 70 percent and the second to 80 percent. In queue 4, it is recommended to set the first threshold to 40 percent and leave the default values for both the second and third thresholds at 100 percent.



Once the queues and thresholds have been defined, traffic can be assigned to queues and thresholds either by CoS values or DSCP values. While DSCP-to-queue and threshold maps override CoS-to-queue and threshold maps, these mappings should be as consistent as possible to ensure predictable behavior and simplify troubleshooting. The following mappings are recommended:

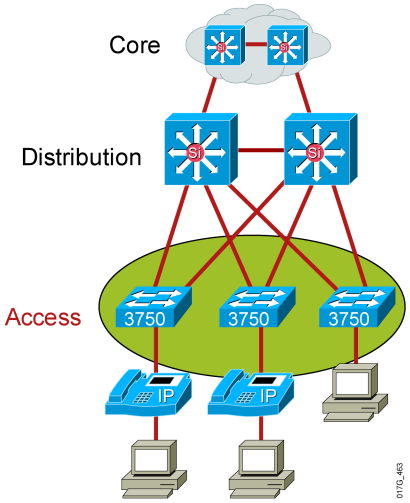
- CoS 0 or DSCP 0 (best-effort traffic) should be mapped to queue 3 threshold 3 (the tail of the queue), because no other traffic is to be assigned to queue 3.
- CoS 1 (scavenger and bulk) should be mapped to queue 4 threshold 3. Scavenger traffic can then be further contained by a DSCP-to-queue and threshold mapping assigning DSCP CS1 to queue 4 threshold 1 (previously set at 40 percent); bulk data using DSCP values AF11, AF12, or AF13 (decimal values 10, 12, and 14, respectively) can then use the remainder of the queue. Bulk data can use either threshold 2 or threshold 3 as its WTD limit (both of which are set to 100 percent).
- CoS 2 and DSCP CS2, AF21, AF22, and AF23 (decimal values 16, 18, 20, and 22, respectively) can be assigned to queue 2 threshold 1 (previously set at 70 percent). This limits network management and transactional data to a subset of queue 2. The temporary marking value for mission-critical traffic, DSCP 25, should also be assigned to queue2 threshold 1.
- CoS 3, along with DSCP CS3 and AF31 (decimal values 24 and 26, respectively) can be assigned to queue 2 threshold 2 (previously set to 80 percent). This allows for preferential treatment of call signaling traffic within queue 2.
- CoS 4 and DSCP CS4, AF41, AF42, and AF43 (decimal values 32, 34, 36, and 38, respectively) can be assigned to queue 2 threshold 1. In this manner, video (both interactive and streaming) does not drown out call signaling or network and internetwork control traffic within queue 2.
- CoS 5 and DSCP EF (decimal value 46) should be assigned to queue 1 threshold 3, because voice is the only traffic to be assigned to the strict-priority queue.

- CoS 6 and DSCP CS6 (decimal value 48) and CoS 7 and DSCP CS7 (decimal value 56) should be assigned to queue 2 threshold 3. In this manner, there is always some room available in queue 2 to service network and internetwork control traffic.

Note The QoS features and configuration syntax are identical for the Catalyst 2970, 3560 and 3750 Series switches.

QoS Design Options for Access-Layer Catalyst 3750 Switches

- Trusted Endpoint Model
- Auto QoS VoIP Model
- Untrusted PC + SoftPhone with Scavenger-Class QoS Model
- Untrusted Server with Scavenger-Class QoS Model
- Conditionally Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model
- Conditionally Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model



The following are QoS design options for access-layer Catalyst 3750 switches:

- **Trusted Endpoint Model:** You can use the **mls qos trust dscp** command in interface configuration mode to configure the switch to trust an endpoint. The trusted endpoint should be assigned either the VVLAN or the data VLAN (DVLAN) with the appropriate switchport commands.
- **Auto QoS VoIP Model:** When you enable AutoQoS VoIP on the Catalyst 3750 switch, the switch automatically generates a QoS configuration based on the traffic type and ingress packet label. If you use the **auto qos voip trust** command, the switch automatically sets the ingress classification to trust the CoS value received in the packet on a nonrouted port by using the **mls qos trust cos** command. If you use the **auto qos voip cisco-phone** command, the switch automatically enables the trusted boundary feature, which uses the Cisco Discovery Protocol to detect the presence or absence of a Cisco IP phone. If you use the **auto qos voip cisco-softphone** command, the switch automatically creates class maps and policy maps. After creating the class maps and policy maps, the switch automatically applies the policy map called AutoQoS-Police-SoftPhone to an ingress interface on which auto-QoS with the Cisco SoftPhone feature is enabled.

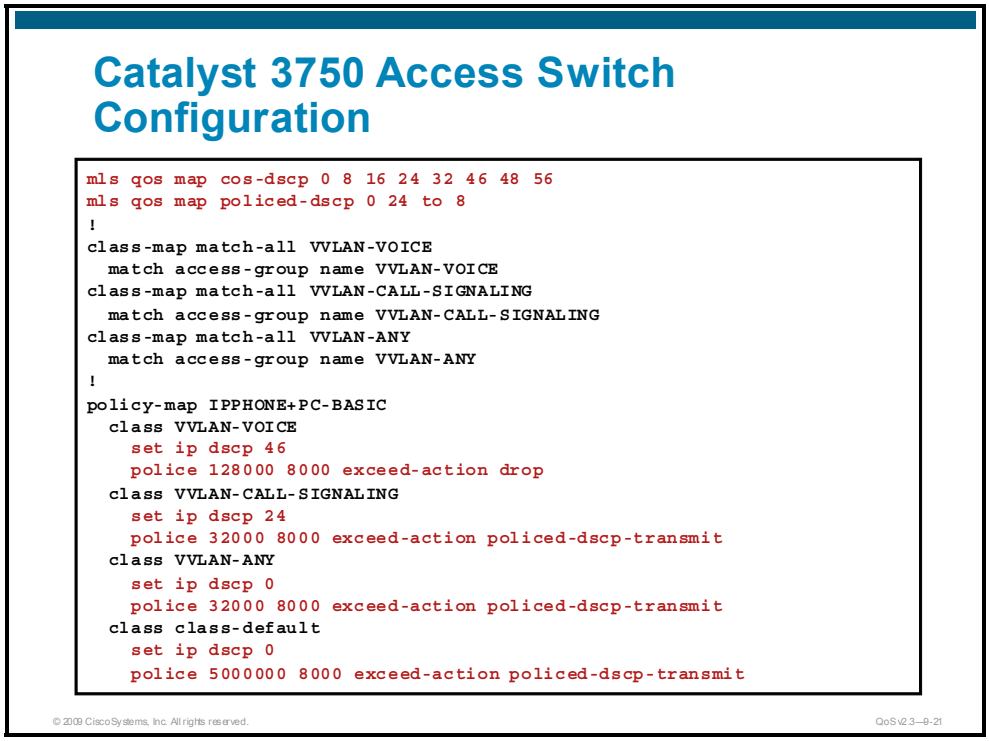
Note	For a list of the commands that the switch applies when you enable AutoQoS VoIP, refer to the “Campus QoS Design” section of the *Enterprise QoS Solution Reference Network Design Guide* on Cisco.com.

- **Untrusted PC Plus Cisco SoftPhone with Scavenger-Class QoS Model:** Configure the switch as follows:
  - Use ACLs to identify Cisco SoftPhone traffic.
  - Mark down out-of-profile Cisco SoftPhone voice traffic, out-of-profile signaling traffic, and out-of-profile data traffic to scavenger (CS1).
  - Remark excess traffic that is marked 0, CS3, or EF to CS1.

- **Untrusted Server with Scavenger-Class QoS Model:** Configure the switch as follows:
  - Remark excess traffic that is marked 0, AF11, AF21, or DSCP 25 to CS1.
  - Mark down out-of-profile excess data traffic to scavenger (CS1).
  - Mark down out-of-profile network application traffic to scavenger (CS1)
- **Conditionally Trusted IP Phone Plus PC with Scavenger-Class QoS (Basic) Model:** Configure the switch as follows:
  - Modify the CoS-to-DSCP mapping to map CoS 5 to DSCP EF.
  - Remark excess VVLAN and DVLAN traffic to scavenger (CS1).
  - Permit only one voice call per switchport VVLAN.
  - Mark down out-of-profile call signaling and out-of-profile data traffic to scavenger (CS1).
  - Mark down unauthorized VVLAN traffic to scavenger (CS1).
  - Configure an ACL to match voice traffic by VVLAN subnet and VoIP UDP port-range.
  - Configure an ACL to match call signaling by VVLAN subnet and call-signaling TCP port-range
  - Configure an ACL to match all other traffic sourced from the VVLAN subnet

Note The Catalyst 3750 switch currently does not fully support per-port per-VLAN policing due to hardware restrictions. For more information on this and other considerations for the Catalyst 3750 switch, refer to the "Campus QoS Design" section of the *Enterprise QoS Solution Reference Network Design Guide* on Cisco.com.

- **Conditionally Trusted IP Phone Plus PC with Scavenger-Class QoS (Advanced) Model:** Configure the switch as follows:
  - Modify CoS-to-DSCP mapping to map CoS 5 to DSCP EF.
  - Remark excess DVLAN traffic that is marked 0, AF11, AF21, CS3, DSCP 25, or AF41 to scavenger (CS1).
  - Permit only one IP/VC stream per switchport.
  - Permit only one voice call per switchport VVLAN.
  - Mark down unauthorized VVLAN traffic to scavenger (CS1).
  - Mark down all call-signaling and data traffic that is out-of profile to scavenger (CS1).
  - Configure an ACL to match voice traffic by VVLAN subnet and DSCP EF.
  - Configure an ACL to match call-signaling traffic by VVLAN subnet call-signaling TCP port-range.
  - Configure ACLs to identify PC applications such as interactive video, mission-critical data, transactional data, and bulk data.
  - Configure an ACL to match all other traffic sourced from the VVLAN subnet.



The figure shows the Conditionally-Trusted IP Phone + PC + Scavenger (Basic) Model configuration on a Catalyst 3750 access switch. The **mls qos map cos-dscp** command modifies CoS-to-DSCP mapping to map CoS 5 to DSCP EF. The **mls qos map policed-dscp** command configures the switch to remark excess VVLAN and DVLAN traffic to scavenger (CS1).

This configuration uses three class maps to classify traffic into three classes (VVLAN-VOICE, VVLAN-CALL-SIGNALING, and VVLAN-ANY) using three extended IP access control lists (ACLs).

The policy map called IPPHONE+PC-BASIC is used to configure the switch to behave as follows:

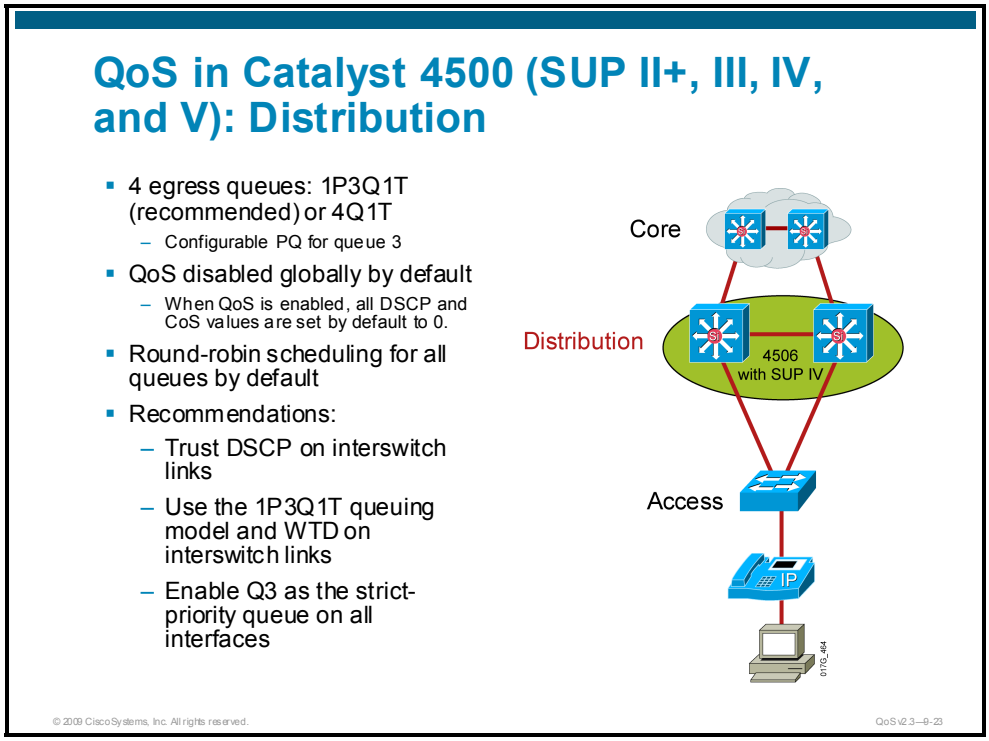
- Permit only one voice call per switchport VVLAN.
- Mark down out-of-profile call signaling to scavenger (CS1).
- Mark down unauthorized VVLAN traffic to scavenger (CS1).
- Mark down out-of-profile data traffic scavenger (CS1)

Catalyst 3750 Access Switch Configuration (Cont.)

```
interface GigabitEthernet0/1
  switchport access vlan 10
  switchport voice vlan 110
  service-policy input IPPHONE+PC-BASIC
!
ip access list extended VVLAN-VOICE
  permit udp 10.1.110.0 0.0.0.255 any range 16384 32767
ip access list extended VVLAN-CALL-SIGNALING
  permit tcp 10.1.110.0 0.0.0.255 any range 2000 2002
ip access list extended VVLAN-ANY
  permit ip 10.1.110.0 0.0.0.255 any
```


The **service-policy** command applies the IPPHONE+PC-BASIC policy to the GigabitEthernet0/1 interface. The three extended access lists identify traffic as follows:

- The ACL named VVLAN-VOICE identifies voice traffic by matching the VVLAN subnet and a VoIP UDP port-range.
- The ACL named VVLAN-CALL-SIGNALING identifies call signaling traffic by matching the VVLAN subnet and a Call-Signaling TCP port-range.
- The ACL named VVLAN-ANY identifies all other traffic sourced from the VVLAN subnet.



Like the Catalyst 3750 switch, the Cisco Catalyst 4500 Series switches with Supervisors II+, III, IV, and V can be found at either the access layer or the distribution layer of the campus. Furthermore, due to their high performance, they may also be found at the core layer of some campus networks. The QoS design recommendations for the distribution layer on the Catalyst 4500 Series switch are as follows:

- Trust DSCP on interswitch links.
- Use the 1P3Q1T queuing model and WTD on interswitch links.

QoS is globally disabled on Catalyst 4500 Series switches by default. The command to enable QoS globally on a Catalyst 4500 Series switch is simply **qos**. When QoS is globally enabled, all DSCP and CoS values are set by default to 0 (which is equivalent to an untrusted state on all ports).

The Catalyst 4500 Series switch supports four egress queues for scheduling, which may be configured in either 4Q1T or 1P3Q1T modes. By default, all queues are scheduled in a round-robin manner. The third transmit queue can be designated as an optional strict-priority queue. This can be enabled by using the **tx-queue** command in interface configuration mode followed by the **priority high** command in interface transmit-queue configuration mode. This queue can be configured to be shaped to a peak limit, such as 30 percent, to allow bandwidth to be available to non-voice applications. This is useful in situations where a trust boundary has been compromised and a DoS or worm attack is saturating voice queues.

This example shows how to configure transmit queue 3 as the high priority queue:

```
switch(config-if)# tx-queue 3
switch(config-if-tx-queue)# priority high
```

Bandwidth allocations can also be assigned to queues (for certain interfaces) using the **tx-queue** command in interface configuration mode followed by the **bandwidth** sub-command. Bandwidth allocations to queues can only be assigned on the following interface types:

- Uplink ports on supervisor engines
- Ports on the WS-X4306-GB line card
- The two 1000BASE-X ports on the WS-X4232-GB-RJ line card
- The first two ports on the WS-X4418-GB line card
- The two 1000BASE-X ports on the WS-X4412-2GB-TX line card

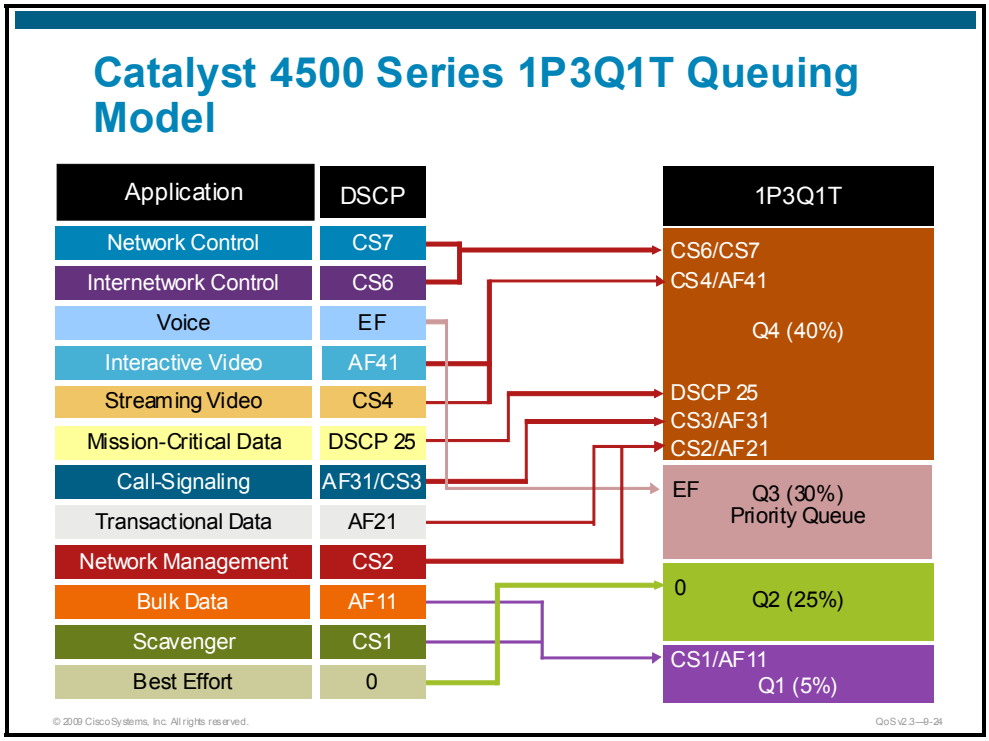
The Cisco Catalyst 4500 Series switches do not support CoS-to-queue mappings, but they do support DSCP-to-queue mappings. DSCP-to-queue mappings can be defined with the **qos map dscp to tx-queue** global command.

While tail-drop or WRED thresholds are not supported on the Catalyst 4500, the switch does support one of the most advanced congestion avoidance mechanisms in the Catalyst family. This congestion avoidance feature is performed by Dynamic Buffer Limiting (DBL). DBL tracks the queue length for each traffic flow in the switch and when the queue length of a flow exceeds its limit, DBL drops packets or sets the (RFC 3168) Explicit Congestion Notification (ECN) bits in the IP packet headers. DBL can be enabled globally with the **qos dbl** global command, as well as on a per-class basis within a policy map with the **dbl** command. A default DBL policy can be applied to all transmit queues.

Given these features and the objective to make queuing consistent across platforms, it is recommended to enable DBL globally on the Catalyst 4500, as well as enable Q3 as the strict-priority queue on all interfaces such that the switch operates in 1P3Q1T mode. This queue can be shaped to 30 percent of the link capacity. Furthermore, Q1 can then be used as the scavenger and bulk queue, Q2 as the best-effort queue, and Q4 as the preferential queue.

On interfaces that support bandwidth allocation, 5 percent could be assigned to Q1, 25 percent to Q2, and 40 percent to Q3. Unlike bandwidth-weights that are used on other platforms, these bandwidth allocations are defined in absolute b/s or as relative percentages of the link bandwidth. In either case, they should not total more than the bandwidth-limit of the link (1 Gb/s or 100 percent), including the priority-bandwidth allocation for Q3.

Note Much of the Catalyst MLS QoS syntax is supported on the Catalyst 4500; however, the **mls** prefix keyword is usually omitted from the configuration commands.



By default, the DSCP-to-queue assignments are as follows:

- DSCP 0-15: queue 1
- DSCP 16-31: queue 2
- DSCP 32-47: queue 3
- DSCP 48-63: queue 4

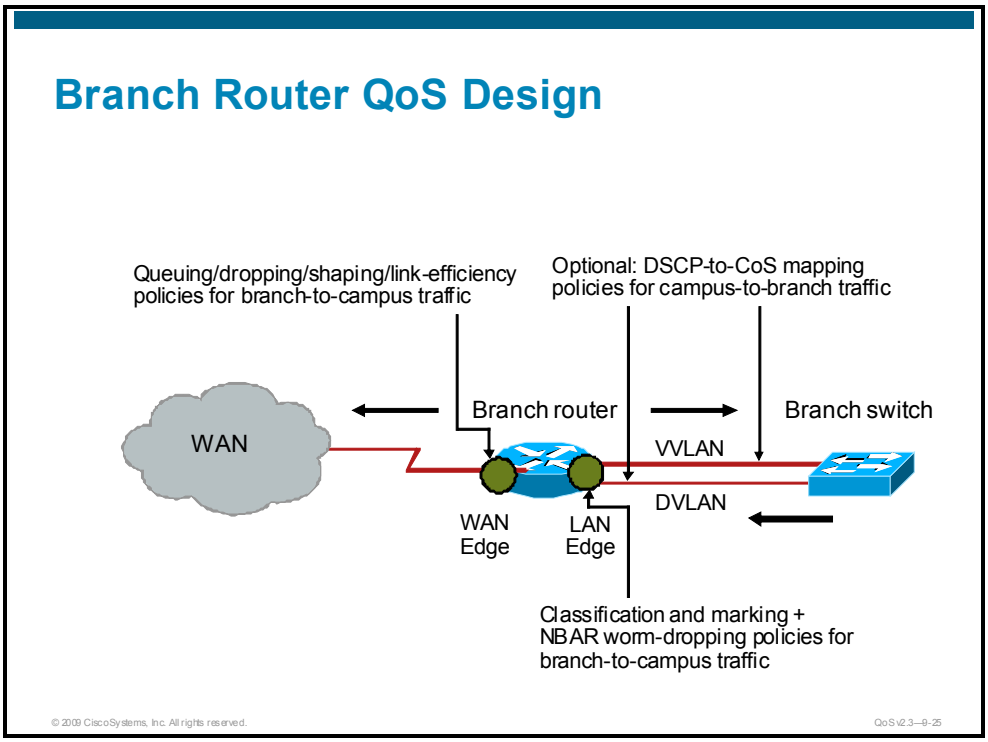
The recommended DSCP-to-queue assignments for the Catalyst 4500 Series switches are as follows:

- DSCP 0: queue 2
- DSCP CS1 (scavenger) and DSCP AF11/AF12/AF13 (bulk data): queue 1
- DSCP CS2 (network management) and AF21/AF22/AF23 (transactional data): queue 4
- DSCP CS3 and AF31 (call-signaling): queue 4
- DSCP 25 (temporary marking for mission-critical data): queue 4
- DSCP CS4 (streaming video) and AF41/AF42/AF43 (interactive video): queue 4
- DSCP EF (Voice): queue 3 (the strict priority queue)
- DSCP CS6 (internetwork control) and CS7 (network control/STP): queue 4

The queuing recommendations for the Catalyst 4500 Series switches (Supervisors II+, III, IV and V) are shown in the figure.

Branch Router QoS Design

This topic describes unique considerations for branch router QoS design.



The following topic discusses the QoS design recommendations for WAN aggregators. For the most part, these designs also apply to branch routers located at the far end of the WAN links. However, at least four unique considerations must be made for branch router QoS design.

One of the first considerations is whether to configure the QoS policies manually or to utilize the Cisco Automatic QoS for the Enterprise (AutoQoS for the Enterprise) feature. AutoQoS for the Enterprise can automatically detect and provision bandwidth for up to 10 classes of traffic. This feature is well suited to smaller branch networks managed by administrators with moderate QoS expertise. However for larger branch networks, where centralized policies are generally preferred, this feature may not be appropriate.

If QoS policies are to be defined manually, other considerations must also be taken into account, such as the presence of unidirectional applications. Some applications, such as streaming video (whether unicast or multicast), require bandwidth allocation only on the WAN edge of the WAN aggregator, not on the WAN edge of the branch router. Therefore, bandwidth allocated to unidirectional applications on the WAN aggregator WAN edge can be redistributed among other preferential classes on the WAN edge of the branch router.

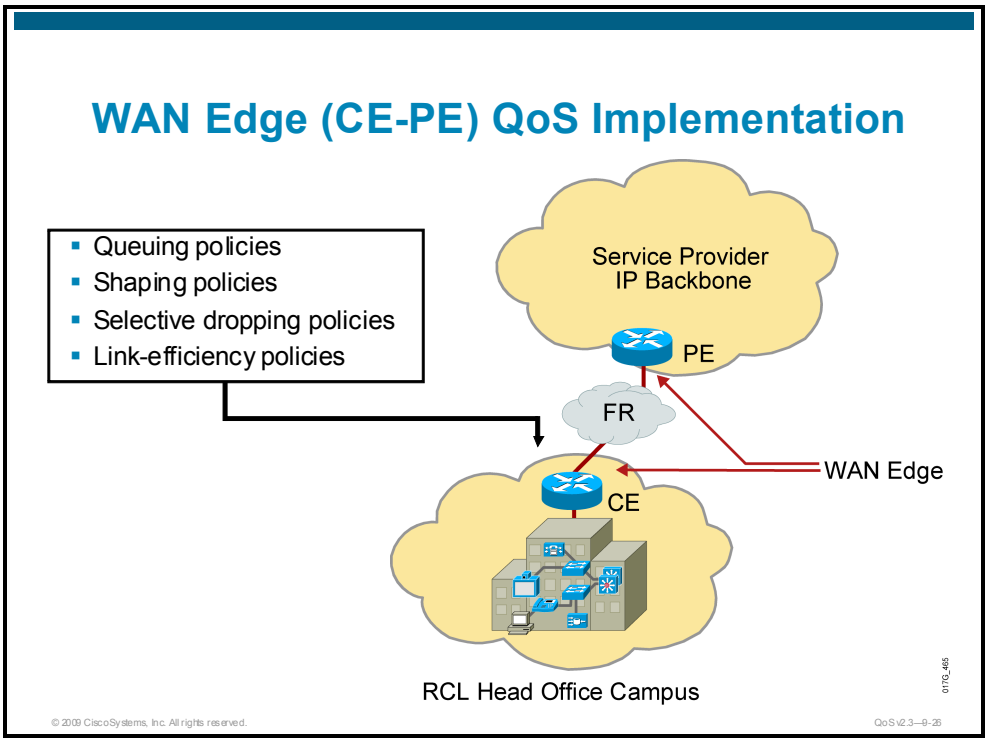
Another characteristic common to branches is that traffic destined to the campus might not be correctly marked on the branch access switches. These switches, which are usually lower-end switches, might or might not have the capabilities to classify by Layer 3 or 4 parameters and mark DSCP values for data applications. Therefore, classification and marking might need to be performed on the LAN edge of the branch router in the ingress direction. Furthermore, branch routers provide the capability to use NBAR to classify and mark flows that require stateful packet inspection.

Related to classification and NBAR, another consideration unique to branch QoS design is that branch routers are a strategic place to deploy NBAR policies for worm identification and policing. NBAR policies can be used to identify and drop Code Red, NIMDA, SQL Slammer, Sasser, and other worms.

The figure shows the QoS policies required on a remote branch router.

WAN Edge QoS Design Considerations

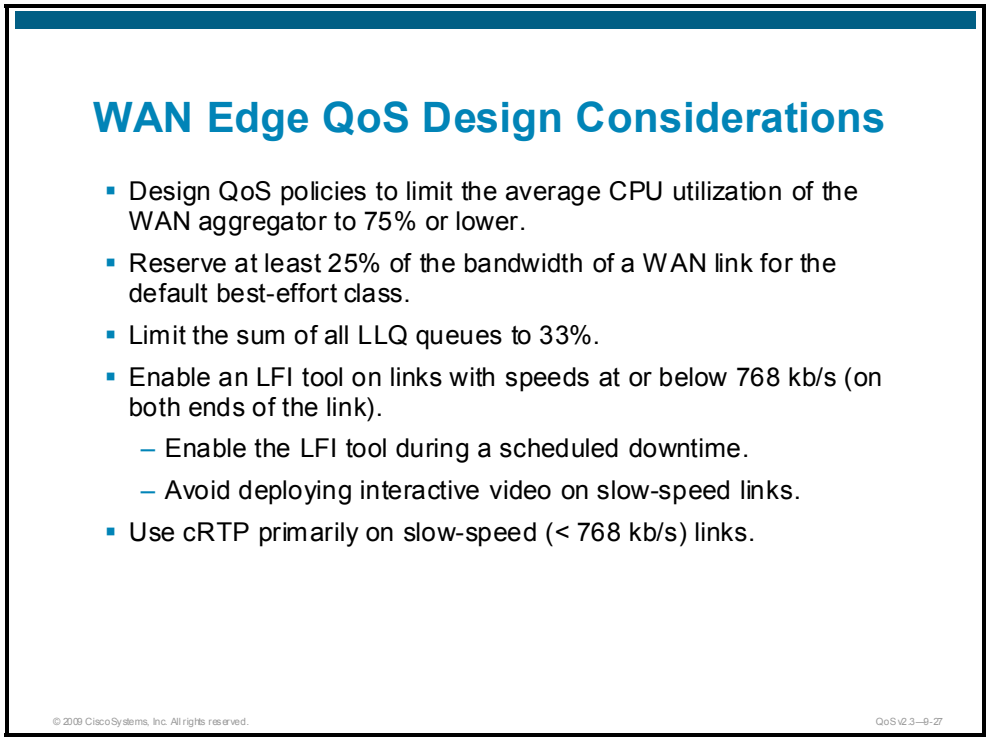
This topic provides design guidance for enabling QoS over the WAN. The recommendations in this topic are not autonomous. They are critically dependent on the recommendations for the enterprise campus discussed in the previous topics.



A fundamental principle of economics states that the more scarce a resource is, the more efficiently it should be managed. In an enterprise network infrastructure, bandwidth is the prime resource and also is the scarcest (and, likewise, most expensive) over the WAN. Therefore, the case for efficient bandwidth optimization using QoS technologies is strongest over the WAN, especially for enterprises that are converging their voice, video, and data networks.

The design principles described in this topic apply primarily to Layer 2 WANs, such as leased lines, Frame Relay, and ATM (including ATM-to-Frame Relay Service Interworking). However, many service providers use these Layer 2 WAN technologies to access Layer 3 VPN services. Therefore, many of the design principles and examples presented in this topic also apply to such VPN access scenarios.

Within typical WAN environments, routers play one of two roles: a WAN aggregator or a branch router. In some very complex WAN models, enterprises might have distributed WAN aggregators to cover regional branches, but the role of such middle-tier routers is not significantly different from that of a WAN aggregator located at a campus edge. This topic focuses on WAN edge recommendations (primarily for WAN aggregator routers), but these correspondingly apply to the WAN edge designs of branch routers. QoS policies required on WAN edges are shown in the figure.



QoS policies required on WAN aggregators include queuing, shaping, selective dropping, and link-efficiency policies in the outbound direction of the WAN link. Traffic is assumed to be correctly classified and marked at Layer 3 before WAN aggregator ingress. Layer 3 markings (preferably DSCP) are media independent and traverse the WAN media, whereas Layer 2 CoS is lost when the media switches from Ethernet to WAN media. Several factors must be kept in mind when designing and deploying QoS polices on WAN edges.

QoS CPU Utilization

WAN edge QoS is performed within Cisco IOS Software. If the WAN aggregator is homing several hundred remote branches, the collective CPU required to administer complex QoS policies might be more than some older devices can provide. The main point is that QoS entails a marginal CPU load. WAN topologies and QoS policies should be designed to limit the average CPU utilization of the WAN aggregator to 75 percent or lower; this allows the router to respond efficiently to routing updates.

Bandwidth Provisioning for Best-Effort Traffic

Reserve at least 25 percent of the bandwidth of a WAN link for the default best-effort class. Because many enterprises have several hundreds, if not thousands, of data applications running over their networks, adequate bandwidth must be provisioned for this class as a whole to handle the sheer volume of applications that default to it.

Bandwidth Provisioning for Real-time Traffic

Limit the sum of all LLQ queues to 33 percent. If too much traffic is assigned to Real-Time (strict-priority/low-latency) queuing, the overall effect is a dampening of QoS functionality for data applications. When real-time applications (such as voice or interactive-video) dominate a WAN link, data applications fluctuate significantly in their response times, destroying the transparency of the "converged" network.

Furthermore, if VoIP traffic is set to dominate a link via low-latency queuing (which is essentially strict-priority FIFO queuing), VoIP actually could negatively impact other VoIP traffic because of extensive FIFO queuing. This easily could result in excessive serialization delays even on medium-speed links (T1/E1 links) where serialization delays ordinarily would not be a consideration. Such excessive serialization delays from VoIP LLQ overprovisioning would increase VoIP jitter and, thus, decrease overall call quality.

Note The 33-percent limit for LLQ queues is a best-practice design recommendation, not a mandate.

Serialization

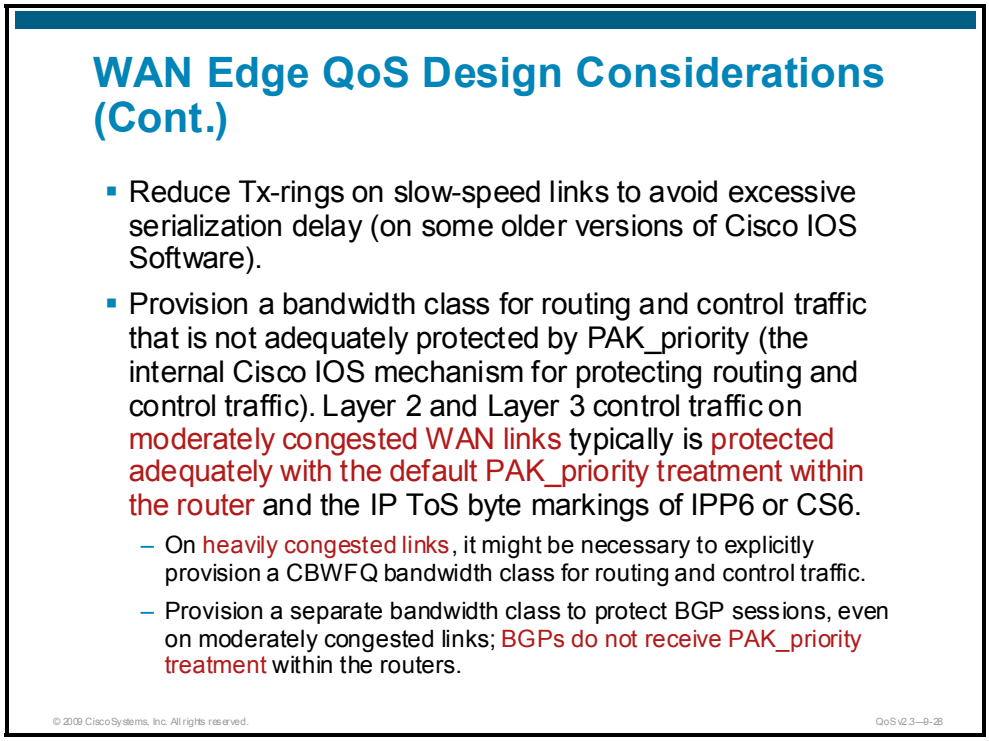
Over the WAN, lower link speeds can cause sufficient serialization delay to adversely affect real-time streams. Serialization delays are variable because they depend not only on the line rate of the link speed, but also on the size of the packet being serialized. Variable delay is also known as jitter. Because the end-to-end, one-way jitter target has been set as 30 ms, the typical per-hop serialization delay target is 10 ms, which allows for up to three intermediate hops per direction of VoIP traffic flow. This 10 ms per-hop target leads to the recommendation that a link fragmentation and interleaving (LFI) tool (either MLP LFI or FRF.12) be enabled on links with speeds at or below 768 kb/s; this is because the serialization delay of a maximum-size Ethernet packet (1500 bytes) takes more than 10 ms to serialize at 768 kb/s and below. Naturally, LFI tools need to be enabled on both ends of the link.

It is recommended that the LFI tool be enabled during a scheduled downtime. Assuming that the network administrator is within the enterprise campus, it is recommended that LFI be enabled on the branch router first (which is on the far end of the WAN link) because this generally takes the WAN link down. The administrator can then enable LFI on the WAN aggregator (the near end of the WAN link), and the link will come back up. Otherwise, if the administrator enables LFI on the WAN aggregator first, the link will go down, along with any in-band management access to the branch router. In such a case, the administrator would need to remove LFI from the WAN aggregator (bringing the link back up), enable LFI on the branch router, and then re-enable LFI on the WAN aggregator.

Additionally, since traffic assigned to LLQ escapes fragmentation, you should not deploy interactive video on slow-speed links; the large interactive video packets (such as 1500-byte full-motion I-frames) could cause serialization delays for smaller interactive video packets.

IP RTP Header Compression Usage

Use cRTP primarily on slow-speed (*less than* 768 kb/s) links. Compressing IP, UDP, and RTP headers (cRTP) for VoIP calls can result in significant bandwidth gains over WAN links. However, cRTP is one of the most CPU-intensive features within the Cisco IOS Software QoS toolset. Therefore, it is recommended that you use cRTP primarily on slow-speed links and carefully monitor CPU usage, especially for WAN aggregators that home a large number of remote branches.



Tx-ring Tuning

Newer versions of Cisco IOS Software automatically size the final interface output buffer (Tx-ring) to optimal lengths for Real-Time applications, such as voice or video. On some older versions of Cisco IOS Software, Tx-rings might need to be reduced on slow-speed links to avoid excessive serialization delay.

To determine the value of the Tx-ring on an interface, use the variation of the **show controllers** command, shown in the following example.

```
switch1#show controllers Serial 1/0 | include tx_limited
tx_underrun_err=0, tx_soft_underrun_err=0, tx_limited=1(64)
```

The value within the parentheses following the **tx_limited** keyword reflects the value of the Tx-ring. In this particular example, the Tx-ring is set to 64 packets. This value can be tuned to the recommended setting of 3 on T1/E1 (or slower) links using the command shown in the following example:

```
switch1(config)#interface Serial 1/0
switch1(config-if)#tx-ring-limit 3
```

The new setting can be verified with the **show controllers** command that is shown above.

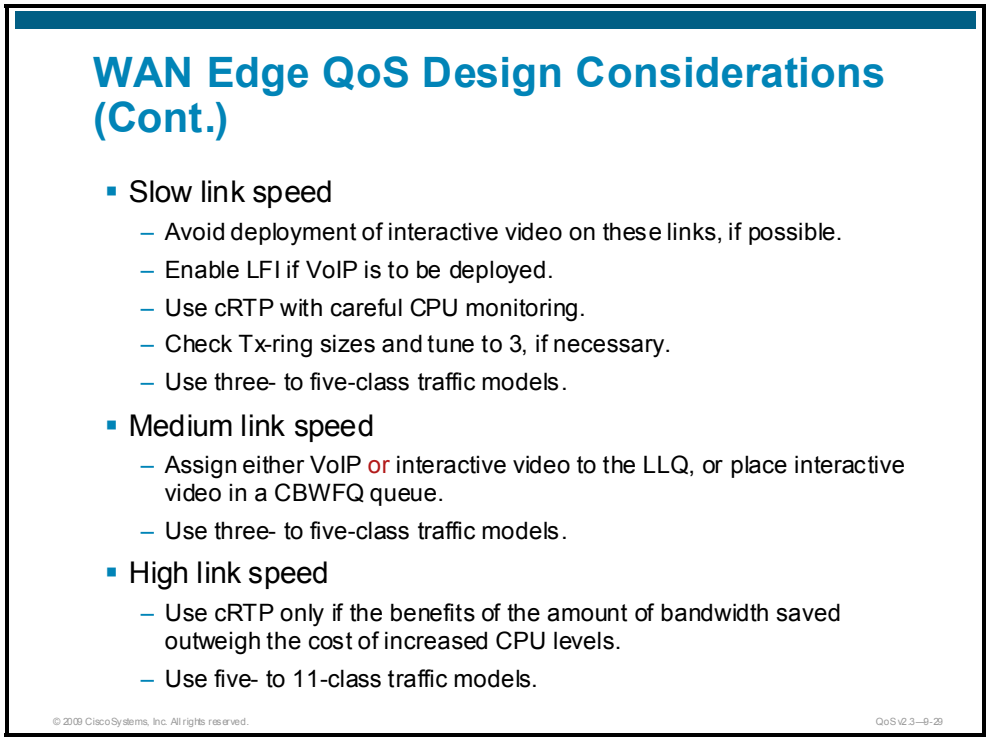
```
switch1#show controllers ser 1/0 | include tx_limited
Tx_underrun_err=0, tx-soft-underru_rr=0, tx-limited=1(3)
```

Note In ATM, the length of the Tx-ring is defined in (576-byte) particles, not packets, and is tuned on a per-PVC basis. On some non-ATM interfaces, the Tx-ring even can be tuned to a minimum of 1 (packet). In either case, the Tx-ring can be tuned (on ≤ 768 kb/s links) to approximately 1500 bytes, which is the MTU of Ethernet.

PAK_Priority

PAK_priority is the internal Cisco IOS Software mechanism for protecting routing and control traffic. The design implications of PAK priority are summarized in the following list:

- Layer 2 and Layer 3 control traffic on moderately congested WAN links typically is protected adequately with the default PAK_priority treatment within the router and the IP ToS byte markings of IPP6 or CS6.
- On heavily congested links, it might be necessary to explicitly provision a CBWFQ bandwidth class for routing and control traffic, as identified by either IPP or CS6.
- Although IS-IS traffic receives PAK_priority within the router, it cannot be marked to IPP6 or CS6 because IS-IS uses a CLNS protocol; it does not use IP. This is important if explicit bandwidth provisioning is required for IS-IS traffic because it cannot be matched against IPP6 or CS6 like most other IGPs. However, NBAR can be used within a class map to match IS-IS traffic.
- Although BGPs (both eBGPs and iBGPs) are marked to IPP6 or CS6, they do not receive PAK_priority treatment within the routers. Therefore, it may be necessary to provision a separate bandwidth class to protect BGP sessions, even on moderately congested links where the underlying IGPs are stable.
- On Catalyst 6500 switches running Cisco IOS Software on both the supervisors and MSFC, IGP packets marked internally with PAK_priority additionally are marked with IPP6 or CS6 and the Layer 2 CoS value of 6. This is because scheduling and congestion avoidance within Cisco Catalyst switches is performed against Layer 2 CoS values.

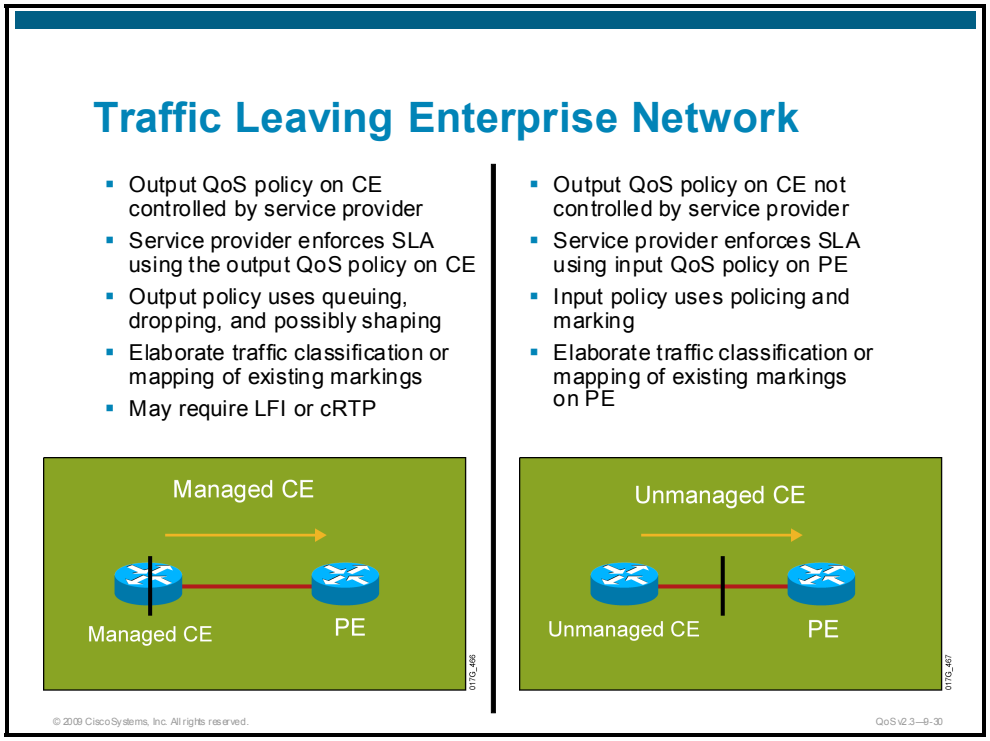


Link Speeds

In the context of WAN links, there are three main groupings of link speeds. These link speeds and their respective design implications are summarized in the following list:

- Slow (link speed ≤ 768 kb/s):
  - Deployment of interactive video generally is not recommended on these links because of serialization implications.
  - These links require LFI to be enabled if VoIP is to be deployed over them.
  - cRTP is recommended (with careful CPU monitoring).
  - Check Tx-ring sizes (especially on slow-speed ATM PVCs); tune to 3, if needed.
  - Three- to five-class traffic models are recommended.
- Medium (768 kb/s ≤ link speed ≤ T1/E1 link speeds):
  - VoIP or interactive video can be assigned to LLQ. There is not typically enough bandwidth to assign both to LLQ and still keep LLQ provisioned at less than 33 percent. Alternatively, interactive video can be placed in a CBWFQ queue.
  - LFI is not required.
  - cRTP is optional.
  - Three- to five-class traffic models are recommended.
- High (T1/E1 link speeds):
  - LFI is not required.
  - cRTP generally is not recommended because the cost of increased CPU levels typically offsets the benefits of the amount of bandwidth saved.
  - Five- to 11-class traffic models are recommended.

It is important to keep in mind that minor differences might exist between QoS configurations on distributed platforms (such as the Cisco 7500 Series Routers with VIPs) and those on nondistributed platforms (such as the 7200 or 1700 Series Routers). The most common difference is the inclusion of the **distributed** keyword after commands such as **ip cef** on distributed platforms.



The QoS requirements on the CE and PE router will differ, depending on whether the CE is managed by the service provider.

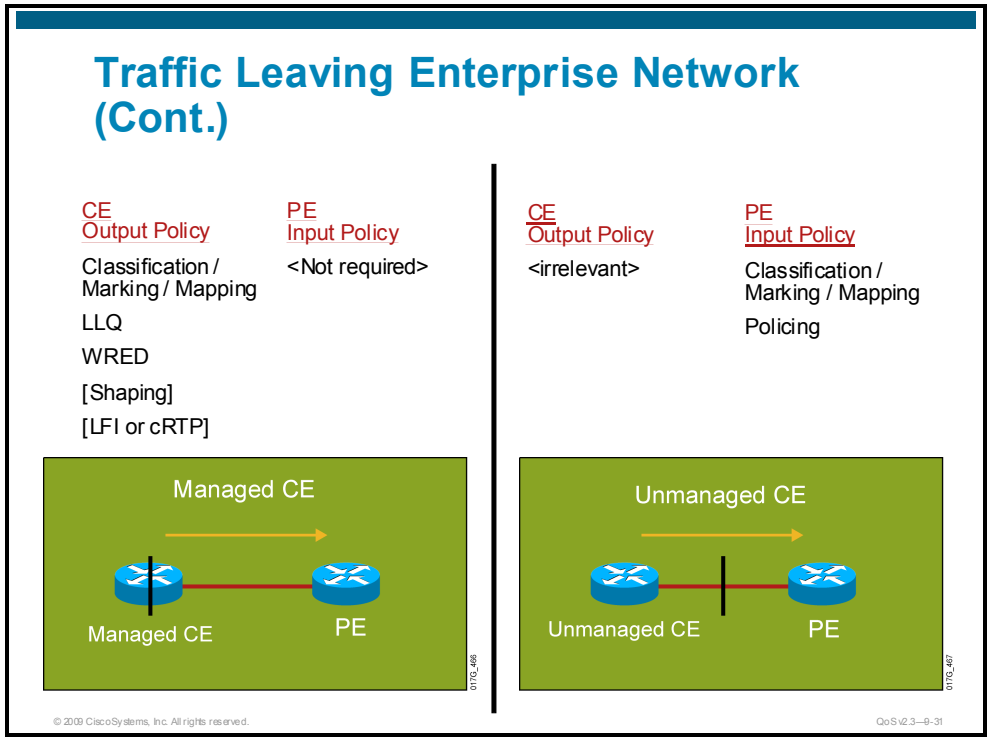
The figure illustrates the general QoS requirements on the CE and PE routers for traffic leaving the enterprise CE router and moving toward the service provider PE router.

For managed CE service, the WAN edge output QoS policy on the CE will be managed and configured by the service provider.

For unmanaged CE service, the WAN edge output QoS policy on the CE will be managed and configured by the enterprise customer.

For managed CE service, the service provider can enforce the SLA for each traffic class using the output QoS policy on the CE. For example, you can use LLQ or CBWFQ to give a maximum bandwidth guarantee to the real-time voice and video traffic class, give a minimum bandwidth guarantee to the data traffic class, and use class-based shaping to provide a maximum rate limit to each data traffic class.

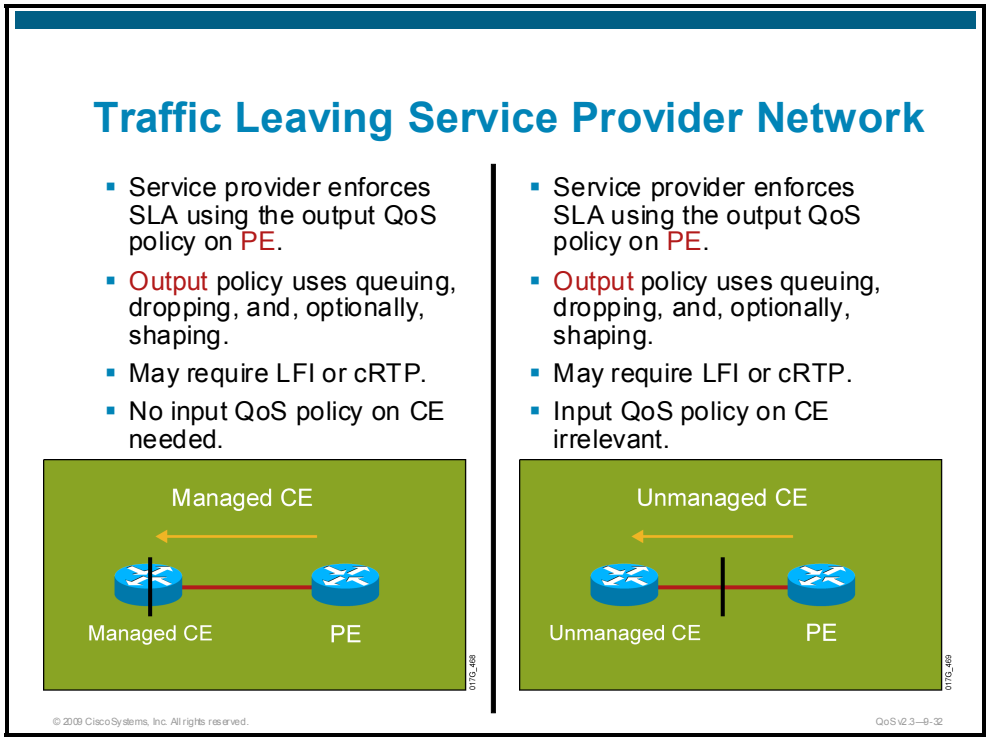
For unmanaged CE service, because the service provider has no control of the CE, the service provider can only enforce the SLA for each traffic class at the input of the PE router. For example, you can use class-based policing to rate-limit the input traffic rate of the different traffic classes and to re-mark the exceeding traffic.



For traffic leaving the CE router, the figure illustrates the different QoS mechanisms that are commonly implemented at the CE and PE routers, depending on whether the CE is managed by the service provider.

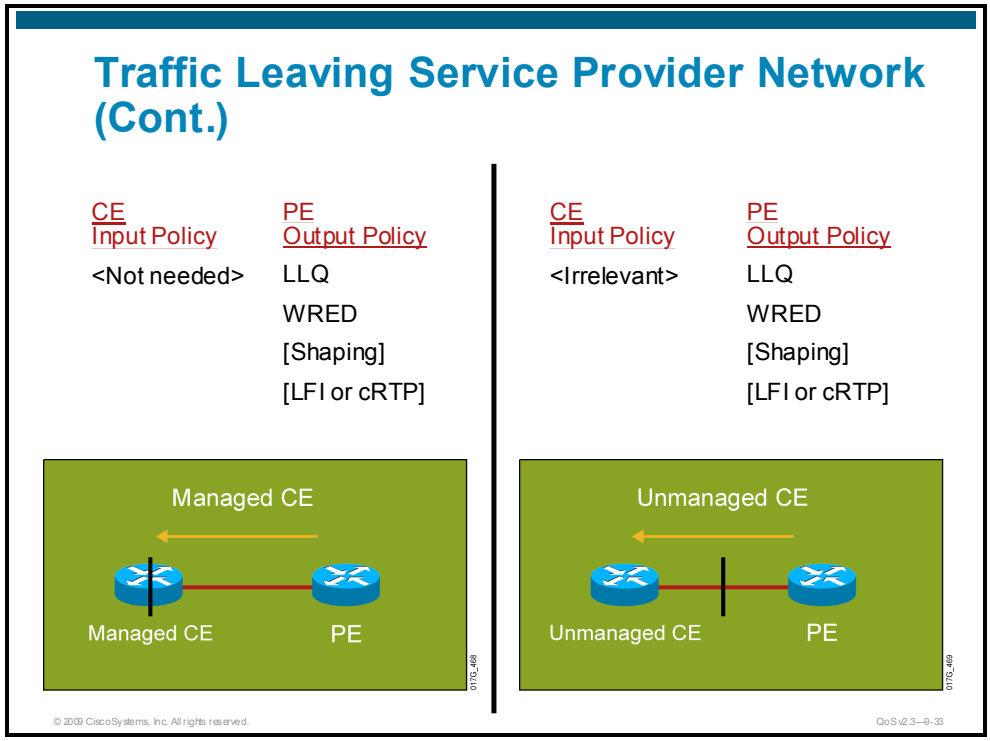
For unmanaged CE, the CE output policy is managed and configured by the enterprise customer; therefore, it is irrelevant to the service provider. At the PE input interface, the service provider will have a policy to classify, mark, or map the traffic. The service provider also typically implements traffic policing to rate-limit the input traffic rate from the enterprise customer, so that the traffic rate does not exceed the contractual rate as specified in the SLA.

For managed CE, the CE output policy is managed and configured by the service provider. The service provider typically has an output policy on the CE router to classify and mark the traffic exiting the CE router. LLQ or CBWFQ (or both) and WRED are used for congestion management and congestion avoidance. To compensate for speed mismatch or oversubscription, traffic shaping may be required. To improve link efficiency, LFI and cRTP are used.



The figure illustrates the general QoS requirements on the CE and PE routers for traffic leaving the service provider PE router toward the enterprise CE router.

For both managed and unmanaged CE service, the service provider can enforce the SLA for each traffic class using the output QoS policy on the PE. For example, use LLQ or CBWFQ to give a maximum bandwidth guarantee to the real-time voice and video traffic class, give a minimum bandwidth guarantee to the data traffic classes, and use class-based shaping to provide a maximum rate limit to each data traffic class.



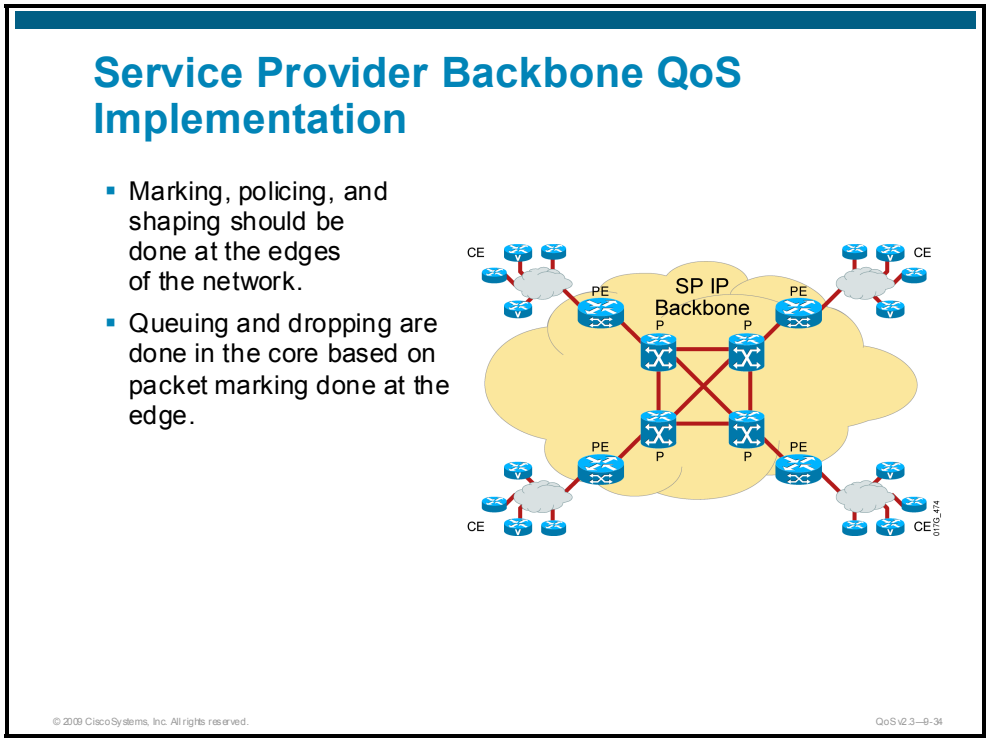
The figure illustrates the different QoS mechanisms that are commonly implemented at the PE router for traffic leaving the service provider PE router toward the enterprise CE router.

For both managed and unmanaged CE service, the service provider typically has an output policy on the PE router using either LLQ or CBWFQ (or both) and WRED, for congestion management and congestion avoidance. To compensate for speed mismatch or oversubscription, traffic shaping may be required. To improve the link efficiency, LFI and cRTP are used.

A customer edge input policy is not required for managed and unmanaged CE services.

Service Provider Backbone QoS Implementations

This topic describes some of the best-practice QoS implementations and configurations on a service provider IP core PE and provider (P) routers.



The figure illustrates the typical QoS configurations required at the PE and P routers within the service provider IP core network.

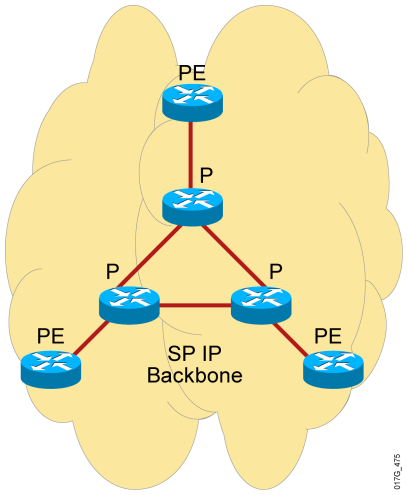
The service provider IP core is used to provide high-speed packet transport. Therefore, all the markings, policing, and shaping should be performed only at the PE router on the PE-to-CE link and not at the core.

Using the DiffServ model, only the edge requires a complex QoS policy. At the core, only queuing and WRED are required. The operation of queuing and WRED will be based on the markings done at the PE.

It is common to implement gigabit switch routers (GSRs) within the service provider IP core. With the GSRs, the queuing mechanism is MDRR. If a router such as the Cisco 7500 Series Router is used in the core, LLQ will be the queuing mechanism.

Service Provider Backbone

- Overprovisioning best-effort backbone is an alternative, but has these drawbacks:
  - Expensive
  - Fate sharing
  - Planning mistakes
  - Failure conditions
  - Unexpected traffic demand
- DiffServ backbone is better



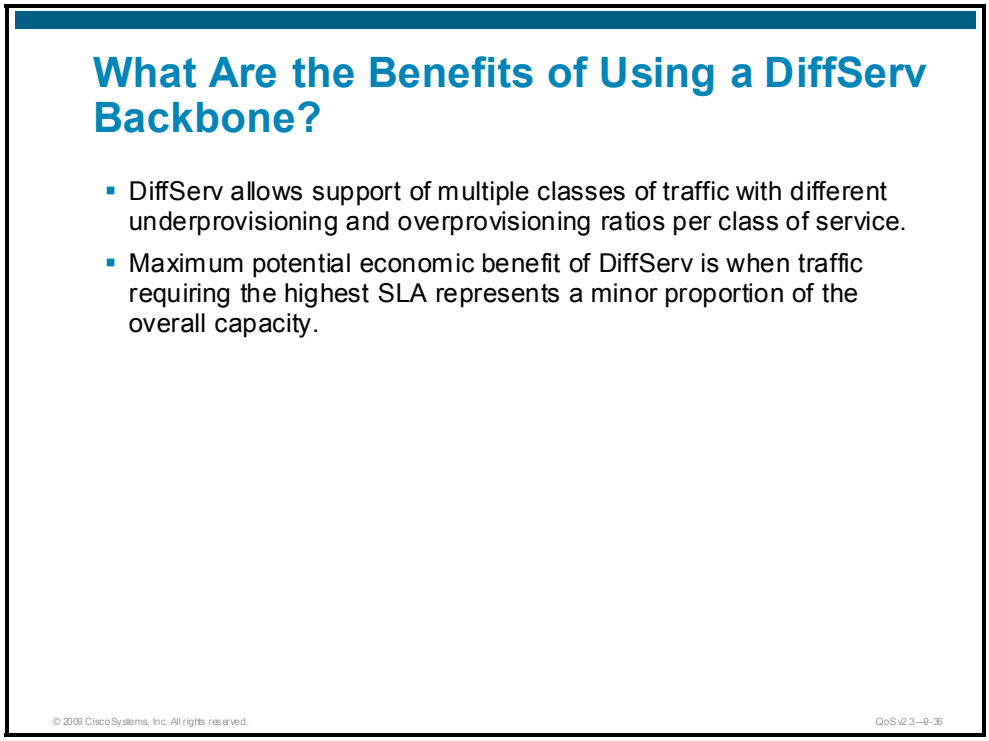
Two of the IP backbone design methods include a best-effort backbone with overprovisioning and a DiffServ backbone.

The more traditional approach is to use a best-effort backbone with overprovisioning. However, to meet the application needs of today (VoIP, videoconferencing, e-learning, and so on), deploying a DiffServ backbone and offering different SLAs for the different traffic classes can greatly reduce the cost and improve the delay, jitter, and packet loss and meet network QoS requirements.

With overprovisioning, a service provider typically uses an overprovisioning factor of 2. For example, if the aggregate traffic load on the network is 10 Gb/s, the network is provisioned for 20 Gb/s of maximum capacity. Some of the problems with a best-effort backbone with overprovisioning include:

- If the capacity planning is not accurate and congestion occurs, because the traffic types are not differentiated, the VoIP packets will not be treated with higher priority than other data traffic, resulting in suboptimal treatment for VoIP packets.
- Overprovisioning for all traffic is very expensive to implement.
- During capacity planning, all of the failure scenarios might not be analyzed. Unplanned failures can cause unexpected congestion in the network. A link or node failure leading to traffic re-routing can take up all the excess capacity.
- The network can experience unexpected traffic demands, which can cause congestions in the network.
- Denial of service attacks on one service will affect all other services.

Using DiffServ, the traffic is isolated into different classes, and each traffic class is provisioned with a different traffic policy based on the QoS requirements of the traffic class. This will reduce the cost and provide better overall latency, delay, and jitter.



The figure lists two of the main benefits of using DiffServ over a best-effort with overprovisioning backbone.

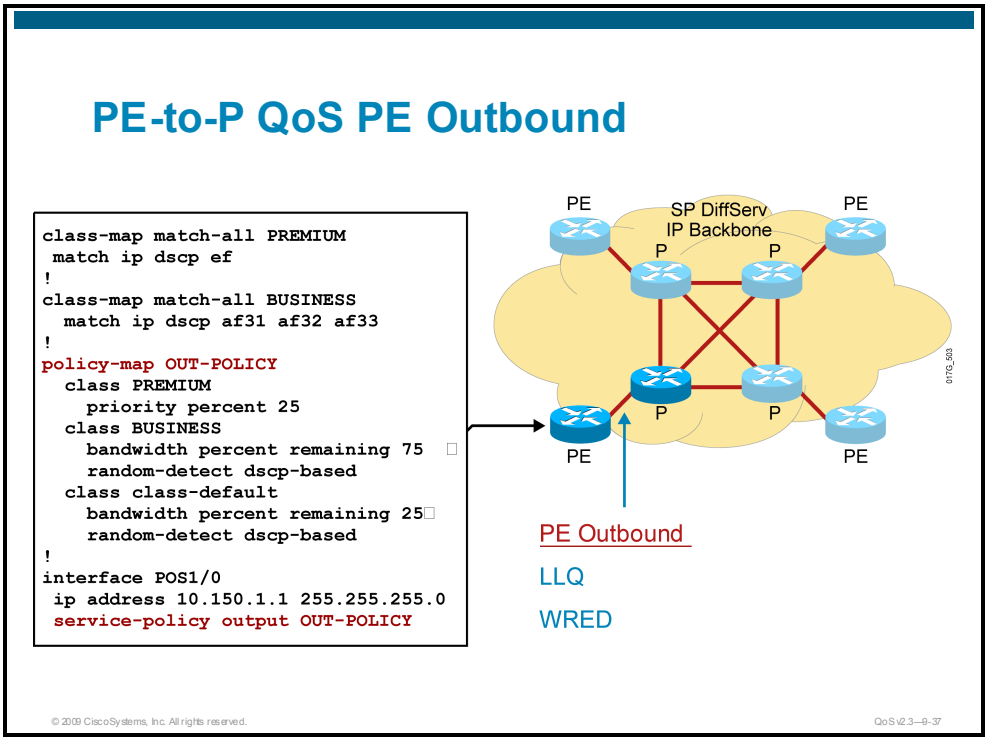
Instead of overprovisioning based on the aggregate bandwidth of all traffic, with DiffServ, each traffic class can be designed with different provisioning ratios. If the premium traffic class is only 20 percent of the total capacity, overprovisioning for 100 percent of the traffic load is expensive and not necessary to guarantee the QoS requirements for the premium traffic class.

The following example illustrates an overprovisioning example with and without DiffServ:

If the total aggregate bandwidth is 10 Gb/s, where the premium class = 2 Gb/s, the business class = 3 Gb/s, and the default class = 5 Gb/s:

- With best effort and an overprovisioning ratio of 2:1, the provisioned bandwidth = 10 Gb/s * 2 = 20 Gb/s.
- With DiffServ, and the premium class having an overprovisioning ratio of 2:1, the business class having a lower overprovisioning ratio of 1.5:1, and the default class not having any overprovisioning, the provisioned bandwidth = (2 Gb/s * 2) + (3 Gb/s * 1.5) + 5 Gb/s = 13.5 Gb/s.

By isolating the traffic into different traffic classes, then treating the different traffic classes with different PHBs, DiffServ can reduce the bandwidth requirement on the network while achieving the same SLA when compared to the non-DiffServ case.



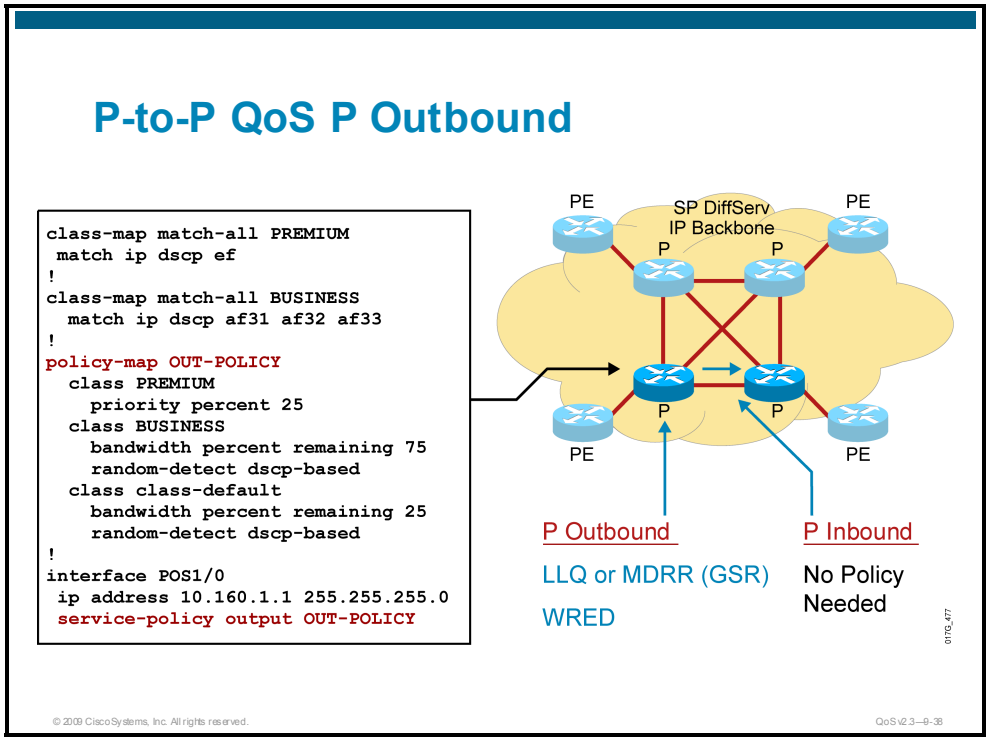
The figure illustrates the typical QoS configurations required at the service provider core PE and P routers.

The complex QoS policies of classification and marking, policing, shaping, LFI, and cRTP are required only at the edge. In the core, only LLQ (or MDRR for GSR) and WRED are needed.

The figure shows the QoS configurations on the ingress PE router outbound interface to the P router.

In this case, a traffic policy called “OUT-POLICY” is configured to provide LLQ or CBWFQ and WRED. Each traffic class bandwidth guarantee is configured using a percentage rather than a fixed bandwidth in kb/s.

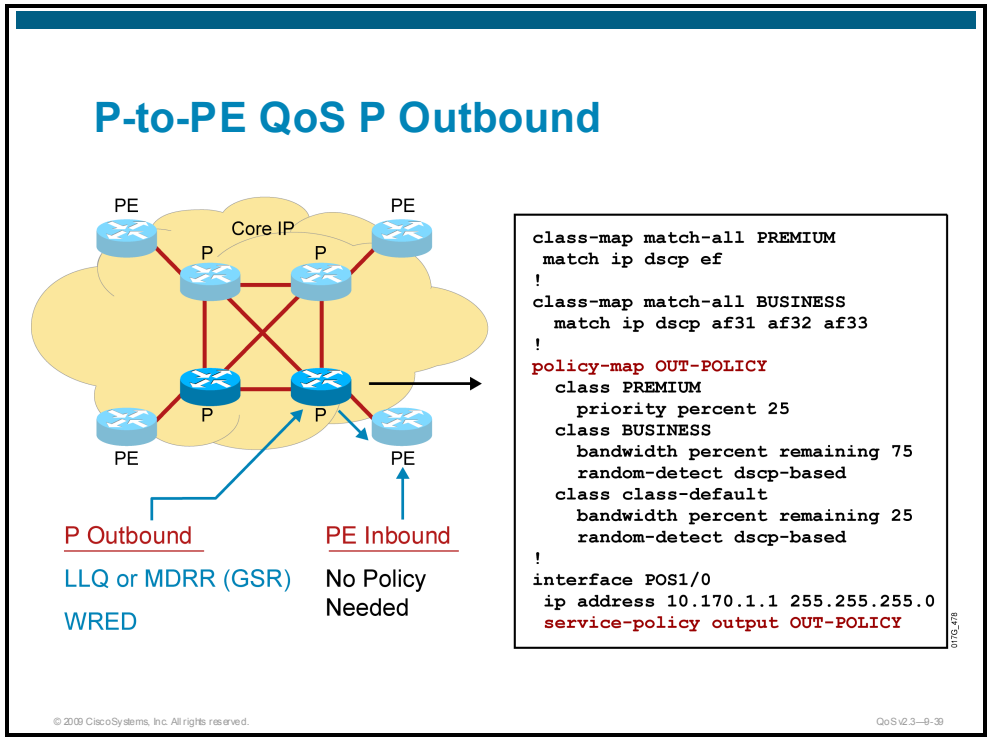
No inbound policy is required on the P router.



The figure shows the QoS configurations on the P router outbound interface to another P router.

In this case, a traffic policy called “OUT-POLICY” is configured to provide LLQ or CBWFQ and WRED. Each traffic class bandwidth guarantee is configured using a percentage rather than a fixed bandwidth in kb/s.

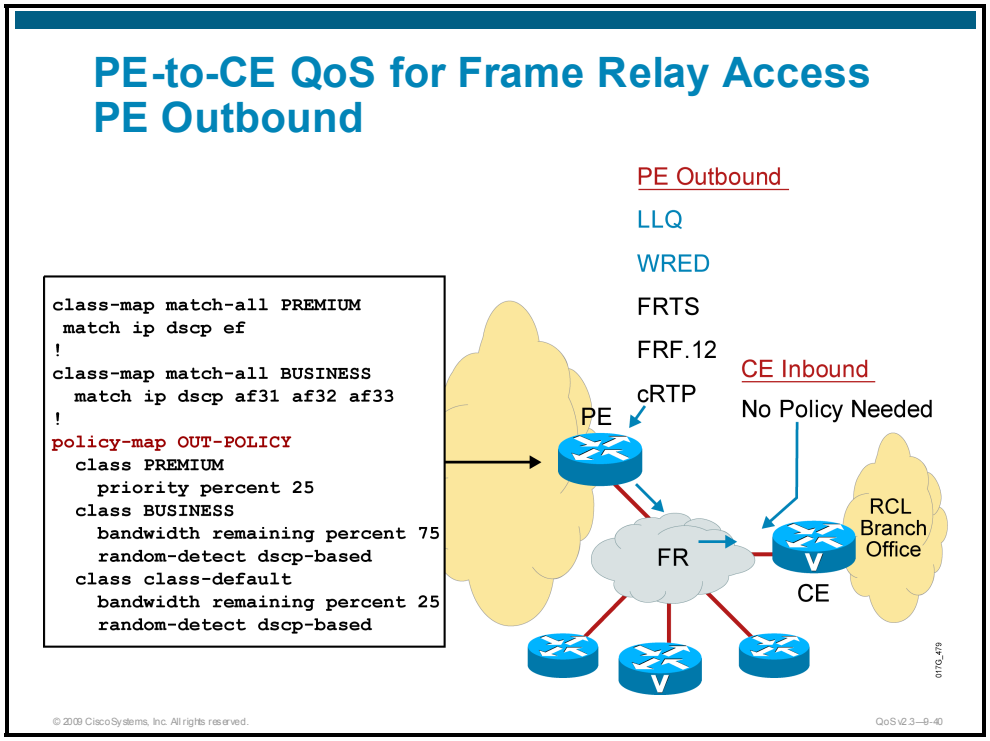
No inbound policy is required on the receiving P router.



The figure shows the QoS configurations on the P router outbound interface to the egress PE router.

In this case, a traffic policy called “OUT-POLICY” is configured to provide LLQ or CBWFQ and WRED. Each traffic class bandwidth guarantee is configured using a percentage rather than a fixed bandwidth in kb/s.

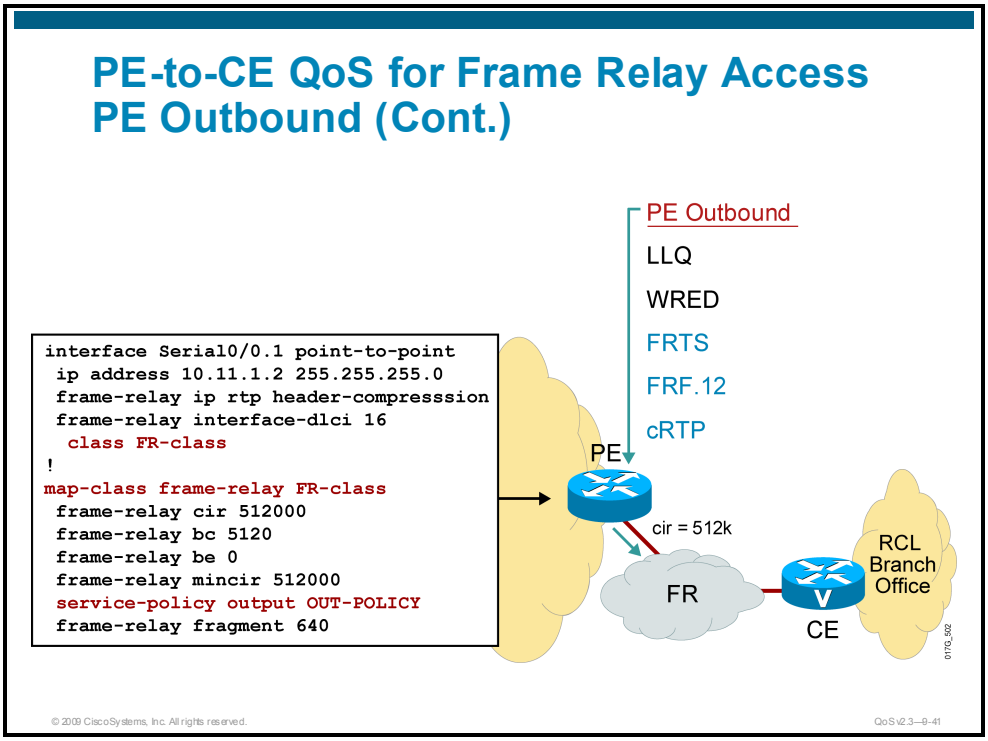
No inbound policy is required on the egress PE router.



The figure shows the QoS configurations on the egress PE router outbound interface to implement the required QoS policy required for each of the three service provider traffic classes.

In this case, a traffic policy called “OUT-POLICY” is configured to provide LLQ or CBWFQ and WRED. Each traffic class bandwidth guarantee is configured using a percentage rather than a fixed bandwidth in kb/s.

No inbound policy is required on the CE router.



In this example, the PE-CE link is a Frame Relay link and FRTS is enabled on the PVC.

FRTS is configured using a Frame Relay map class with a CIR of 512 kb/s, a Bc of 5120 bits, a Be of 0 (no bursting), and a minCIR of 512 kb/s. The CIR is the rate at which you want to normally send when there is no congestion. The CIR needs to be the remote end-link speed or the actual CIR on a virtual circuit. The Bc is the amount you will send per time interval. The CIR and Bc will be used to compute a Tc, where Tc = Bc / CIR. For FRTS, CLI will only allow Bc values that would result in a 125 ms > Tc >10 ms. A recommended value for Tc is 10 ms. To get a Tc of 10 ms, the Bc should set to 1/100 of the CIR.

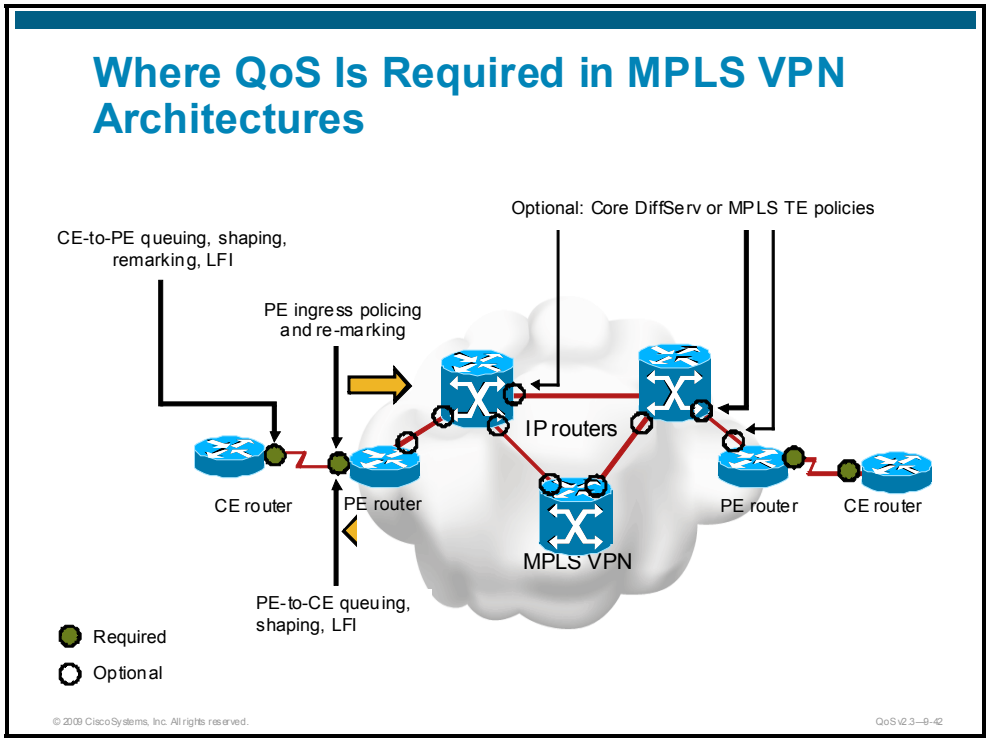
FRF.12 fragmentation and interleaving and cRTP are also enabled within the Frame Relay map class. The fragment size in bytes is set to derive a maximum delay of 10 ms to 15 ms. The fragment size should be the same on both ends.

The “OUT-POLICY” traffic policy is applied within the Frame Relay map class.

Note	If the FR PVC is carrying VoIP traffic, the recommendation is to set the minCIR = CIR and not use adaptive traffic shaping, because by the time the BECNs are received by the sender, drops are already happening on the Frame Relay switch.

MPLS VPN QoS Design

This topic presents QoS design principles and designs to achieve end-to-end service levels over MPLS VPNs.



As noted in the previous lesson, MPLS VPNs are rapidly gaining popularity as private WAN alternatives. 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, enterprise and service-provider QoS designs must be consistent and complimentary.

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 need to co-manage QoS in a cooperative and complementary fashion to achieve end-to-end service levels.

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 WAN aggregator router. The WAN aggregator controls not only campus-to-branch traffic, but also branch-to-branch traffic. Under traditional hub-and-spoke models, QoS is principally administered by the enterprise customer.

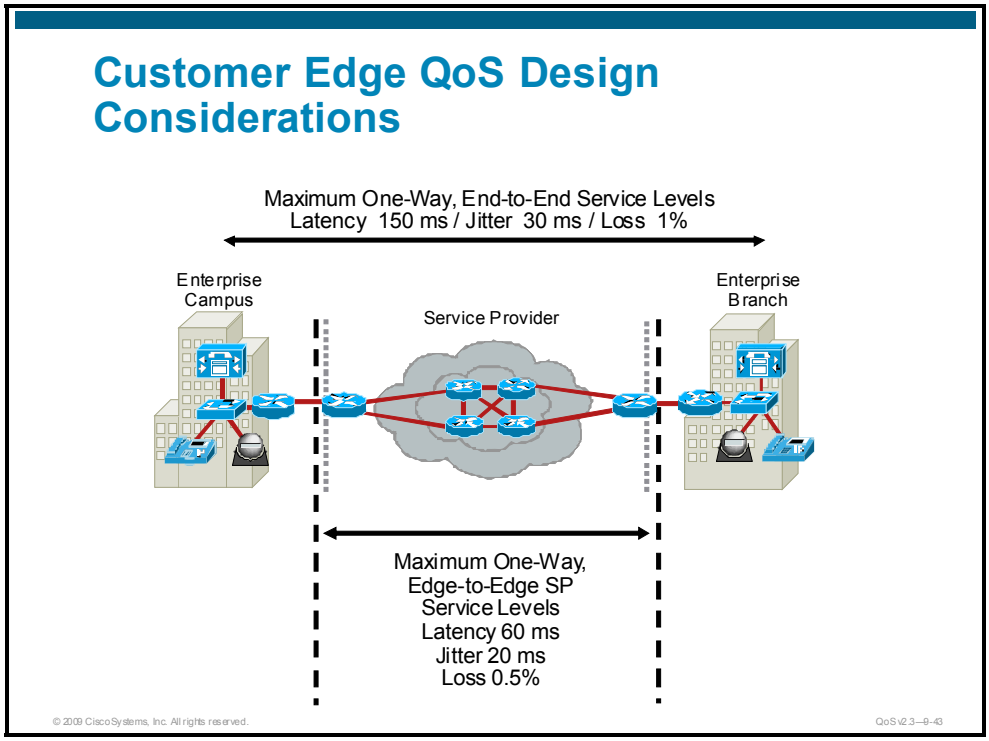
With the advent of MPLS VPN service offerings that inherently offer full-mesh connectivity, the QoS administration paradigm shifts. Under a full-mesh design, the hub router still administers QoS for all campus-to-branch traffic, but it no longer fully controls the QoS for branch-to-branch traffic. Although 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 IP/VC. As with a traditional Layer 2 WAN design, a scheduling policy to prioritize IP/VC on the WAN aggregator is required. The enterprise must also properly provision similar priority scheduling for IP/VC on the branch routers. In this manner, any IP/VC 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 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.

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. The figure summarizes the points where QoS policies can be provisioned within MPLS VPN architectures.



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

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. 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.

In addition to FR and ATM for access, some service providers support Ethernet and 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 lesson.

Service Provider Service-Level Agreements

End-to-end QoS is like a chain that is only as strong as the weakest link. Therefore, it is 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 component of the SLA must be considerably tighter. These SLAs are defined for Cisco-Powered Networks-IP Multiservice Service Providers:

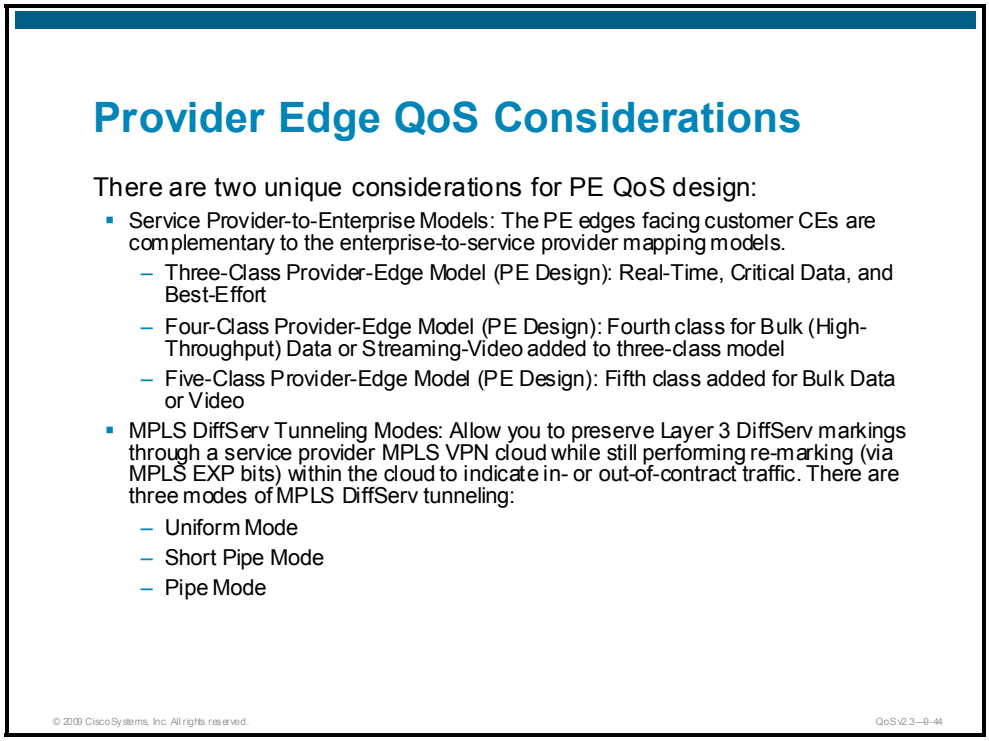
- No more than 60 ms of one-way latency from edge to edge
- No more than 20 ms of jitter
- No more than 0.5 percent loss

The figure illustrates the interrelationship of these SLAs.

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.

Enterprise-to-Service Provider Mapping Models

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 the QoS models of their service providers. For information on how best to collapse and integrate enterprise classes into various service-provider QoS models, refer to Module 9, Lesson 1.



PE designs are relevant for service providers and for enterprises that are self-managing their own MPLS VPNs. There are two unique considerations for PE QoS design:

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

Service Provider-to-Enterprise Models

The PE edges facing customer CEs are complementary to the enterprise-to-service provider mapping models discussed previously.

- **Three-Class Provider-Edge Model (PE 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 (for customer routing traffic), AF31, or CS3. All other code points are re-marked to 0 by an ingress policer. 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.
- **Four-Class Provider-Edge Model (PE Design):** Building on the previous model, 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.
- **Five-Class Provider-Edge Model (PE Design):** Building again on the previous model, a fifth class is added that can be used for either bulk data or video (whichever was not used under the four-class model).

Note For more information on the provider edge mapping models, refer to Module 9, Lesson 1 of this course and to the “MPLS VPN QoS Design” section of the *Enterprise QoS Solution Reference Network Design Guide* on Cisco.com.

MPLS DiffServ Tunneling Modes

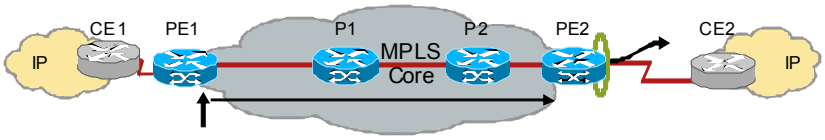
Some service providers re-mark packets at Layer 3 to indicate whether traffic is in contract or out-of-contract. Although this conforms to DiffServ standards, such as RFC 2597, this is not always desirable from the standpoint of the enterprise customer. Because MPLS labels include 3 bits that commonly are used for QoS marking, it is possible to tunnel DiffServ, that is, preserve Layer 3 DiffServ markings through a service provider 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 further in the following pages:

- Uniform Mode
- Short Pipe Mode
- Pipe Mode

MPLS DiffServ Tunneling Modes

Three modes of interaction defined between markings:

- Uniform mode:
  - EXP value is changed in the provider core. At the egress PE, subscriber DSCP/ToS field values are altered. Subscriber will have to reset the original value on the CE device.
- Pipe mode:
  - Provider uses own EXP values including egress-PE-CE link but does not alter subscriber DSCP/ToS values. Subscribers receive traffic with their original DSCP/ToS marked values.
- Short pipe mode:
  - Provider changes EXP values in the core, but honors subscriber DSCP/ToS values on the egress-PE-CE link. Subscribers receive traffic marked with the original DSCP/ToS value.

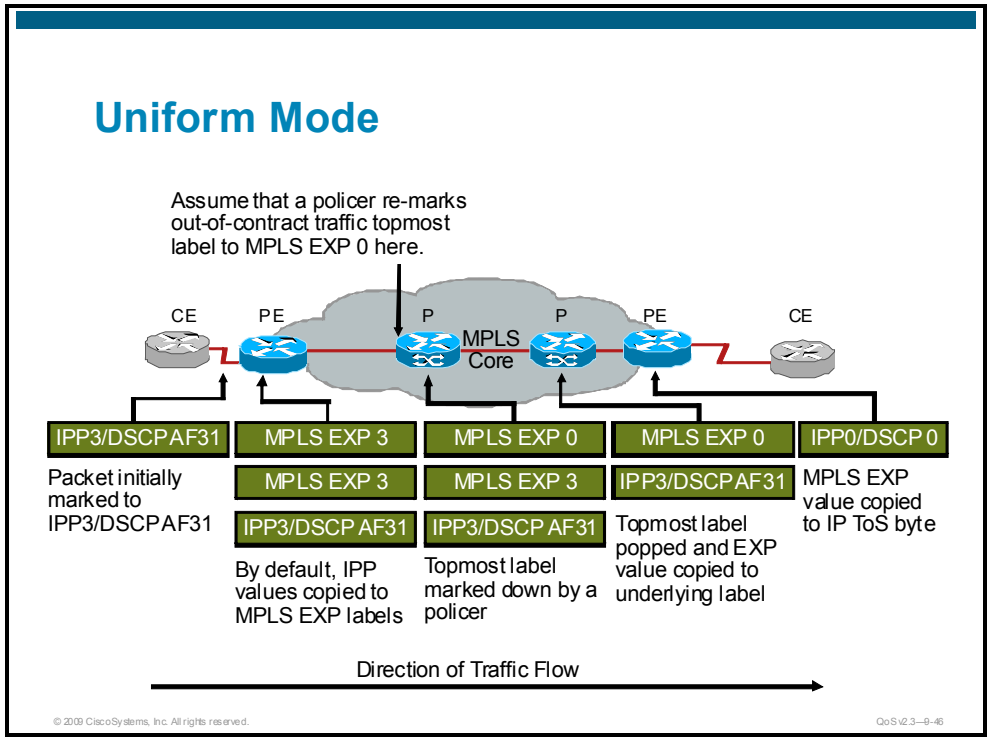


As described in previous examples, some service providers re-mark packets at Layer 3 to indicate whether traffic is in contract or out of contract. Although this conforms to DiffServ standards, such as RFC 2597, this is not always desirable from the standpoint of an enterprise customer.

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 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 the following modes of MPLS DiffServ tunneling:

- Uniform mode
- Short pipe mode
- Pipe 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.

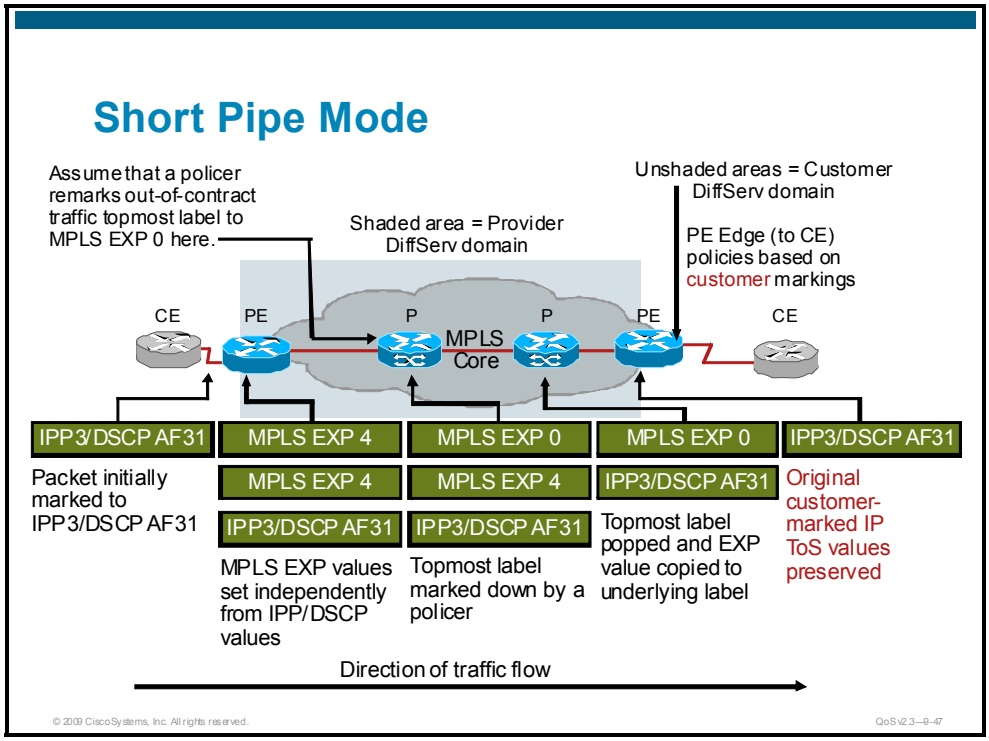
In uniform mode, which is the default mode, the first 3 bits of the IP ToS field (IP Precedence bits) automatically are mapped to the MPLS EXP bits on the ingress PE as labels are pushed onto the packets.

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). The figure shows the behavior of uniform mode MPLS DiffServ tunneling.

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 Software 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).

Note For more information on MPLS VPNs, please refer to the “MPLS VPN QoS Design” section of the *Enterprise QoS Solution Reference Network Design Guide* on Cisco.com.



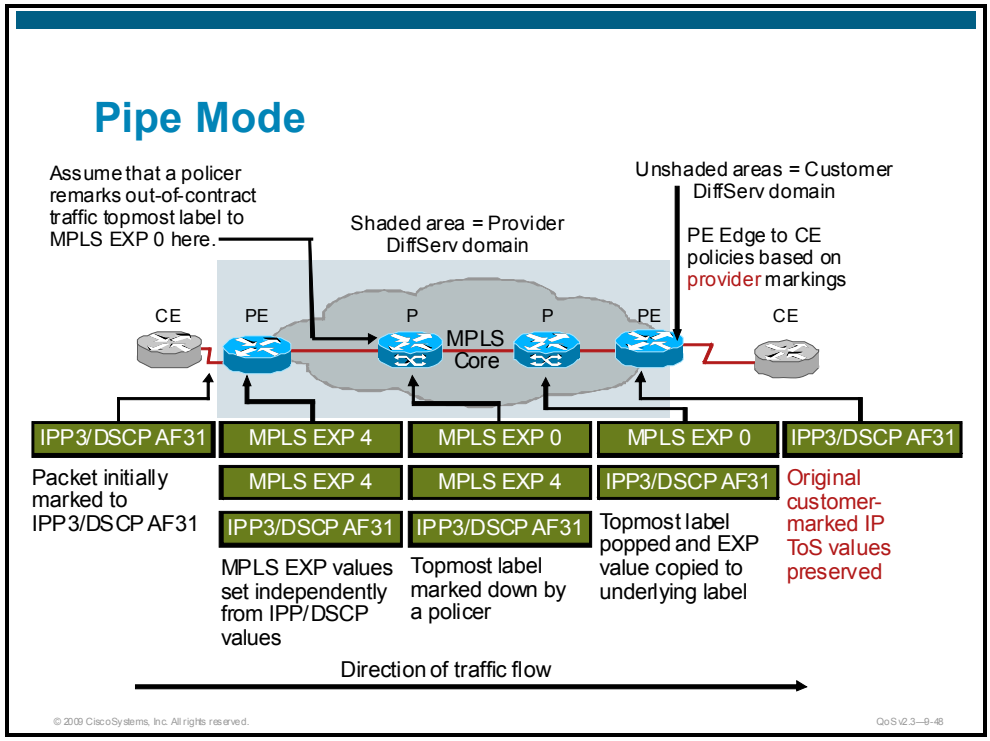
Short-pipe mode is utilized when the customer and service provider are in different DiffServ domains. (The DiffServ domain of the service provider begins at the ingress interface of the ingress PE and terminates on the ingress interface of the egress PE.)

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.

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

In the case of any re-marking occurrence within the service provider MPLS VPN cloud, changes are limited to MPLS EXP re-marking only and are not propagated down to the ToS byte of the underlying IP packet. The figure shows the operation of short-pipe mode MPLS DiffServ tunneling.

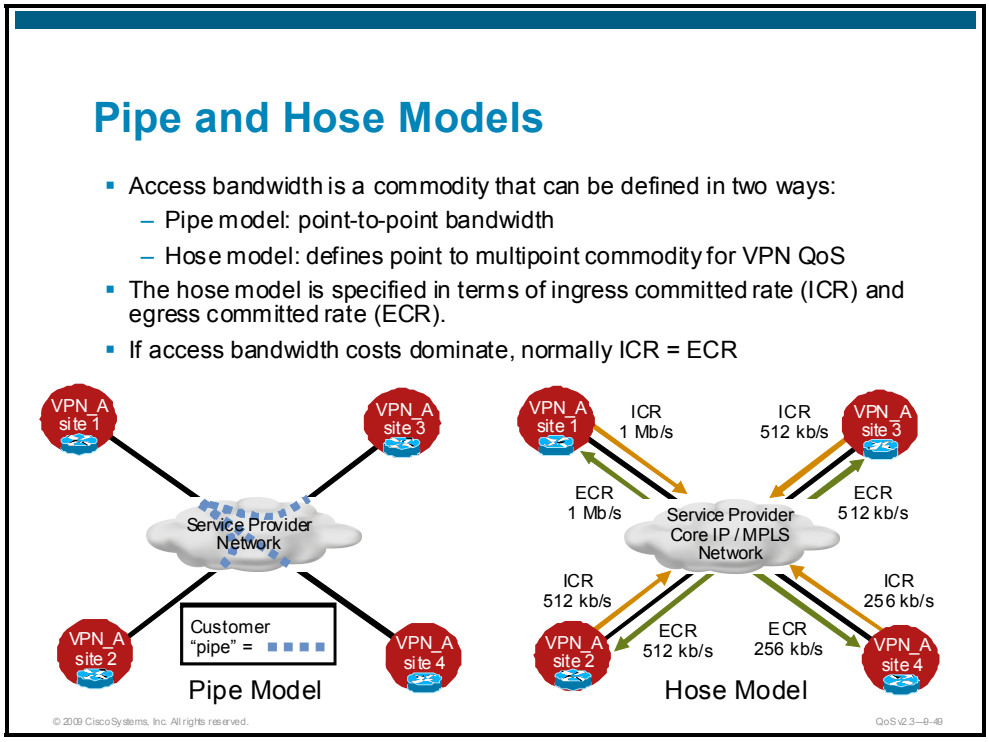
MPLS EXP values can be marked in any way that the provider wants to provide local significance.



Like short-pipe mode, pipe mode provides DiffServ transparency through the service provider network. 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 explicit markings and re-markings of the service provider, not the IP DiffServ markings of the enterprise customer (although these are preserved). As with short-pipe mode, any changes to label markings that occur within the service provider 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 and discard class values. The figure illustrates the pipe mode MPLS DiffServ 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.



There are two models used in the context of MPLS VPNs that refer to how QoS guarantees are offered.

- Point-to-point (pipe model)
- Point-to-cloud (hose model)

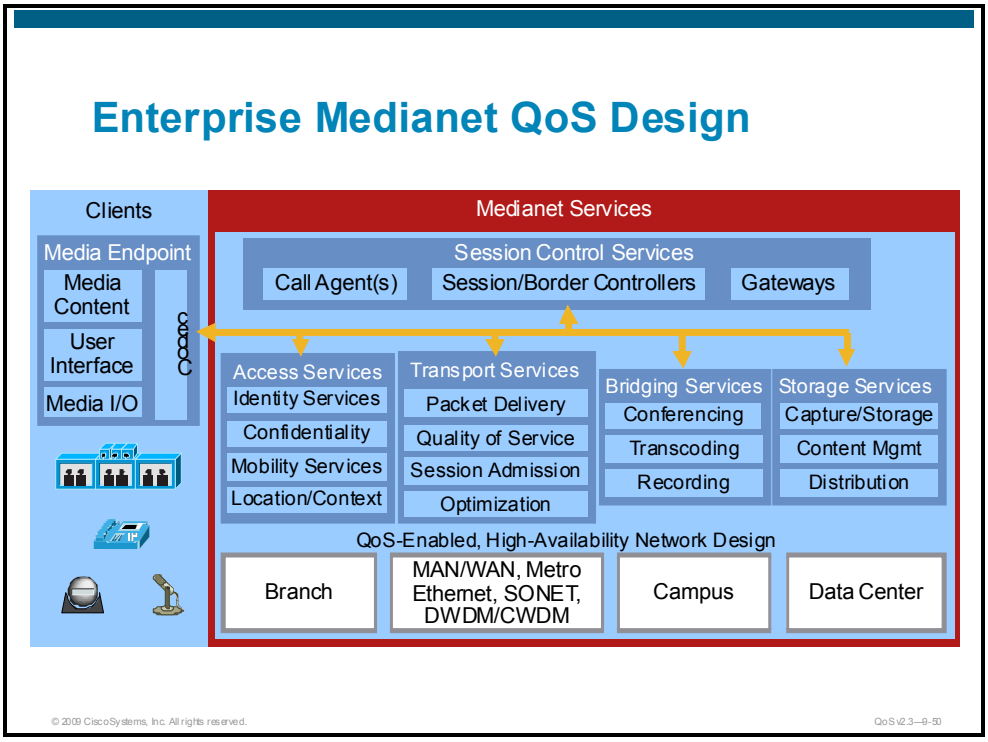
In the point-to-point model, VPN sites have specific QoS guarantees to other VPN sites. In a VPN with three sites, each site will have specific QoS guarantees to the other two sites that belong to the VPN. This approach can be used to offer hard QoS guarantees equivalent to the guarantees available with 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.

QoS Recommendation Summary

This topic summarizes the recommendations for end-to-end QoS.



As media applications increase on the IP network, QoS will play a progressively vital role to ensure the required service level guarantees to each set of media applications, all without causing interference to each other. Therefore, the QoS strategies must be consistent at each place in the network, including the campus, data center, branch WAN/MAN/VPN, and branch.

Integration will play a key role in two ways. First, media streams and endpoints will be increasingly leveraged by multiple applications. For example, desktop video endpoints may be leveraged for desktop video conferencing, web conferencing, and for viewing stored streaming video for training and executive communications.

Additionally, many media applications will require common sets of functions, such as transcoding, recording, and content management. To avoid duplication of resources and higher implementation costs, common media services need to be integrated into the IP network so they can be leveraged by multiple media applications.

Furthermore, because of the effectiveness of multimedia communication and collaboration, the security of media endpoints and communication streams becomes an important part of the media-ready strategy. Access controls for endpoints and users, encryption of streams, and securing content files stored in the data center are all part of a required comprehensive media application security strategy.

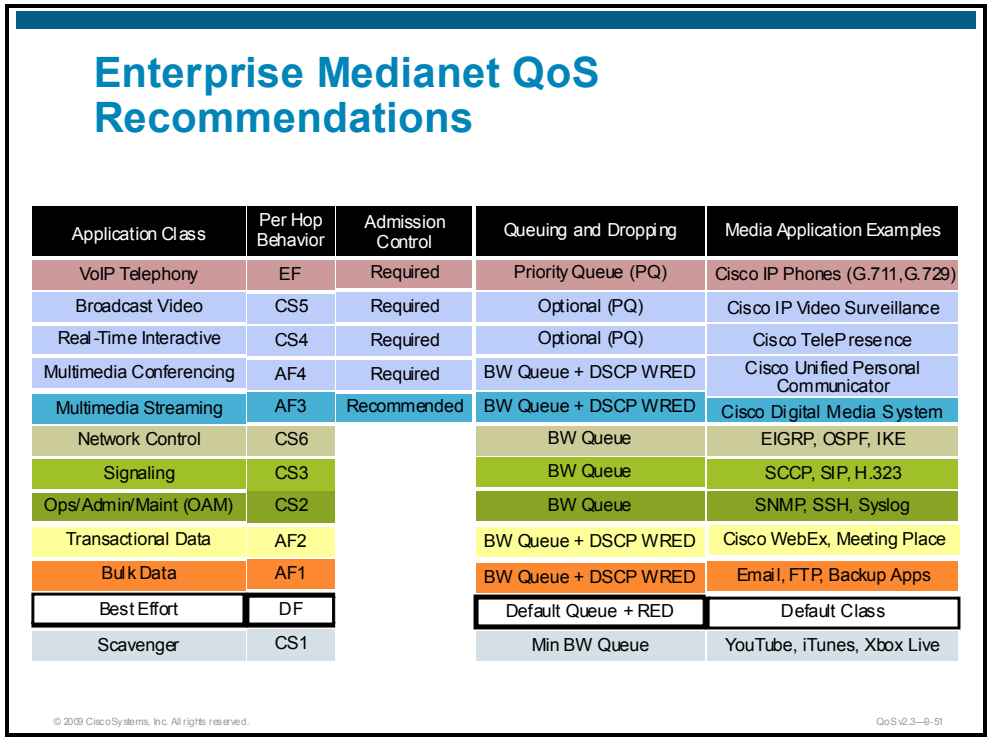
Finally, as the level of corporate intellectual property migrates into stored and interactive media, it is critical to have a strategy to manage the media content, setting and enforcing clear policies, and having the ability to protect intellectual property in secure and managed systems. Just as companies have policies and processes for handling intellectual property in document form, they also must develop and update these policies and procedures for intellectual property in media formats.

Therefore, to meet all these media application requirements, Cisco recommends a medianet architecture. A medianet is built upon an architecture that supports the different models of media applications and optimizes their delivery, such as those shown in the architectural framework in the figure.

An enterprise medianet framework starts with and end-to-end QoS-enabled network infrastructure designed and built to achieve high availability, including the data center, campus, WAN, and branch office networks. The network provides a set of services to video applications, including:

- **Access services:** Provide access control and identity of video clients, as well as mobility and location services.
- **Transport services:** Provide packet delivery, ensuring the service levels with QoS and delivery optimization.
- **Bridging services:** Provide transcoding, conferencing, and recording services.
- **Storage services:** Provide content capture, storage, retrieval, distribution, and management services.
- **Session control services:** Provide signaling and control to setup and tear-down sessions, as well as gateways.

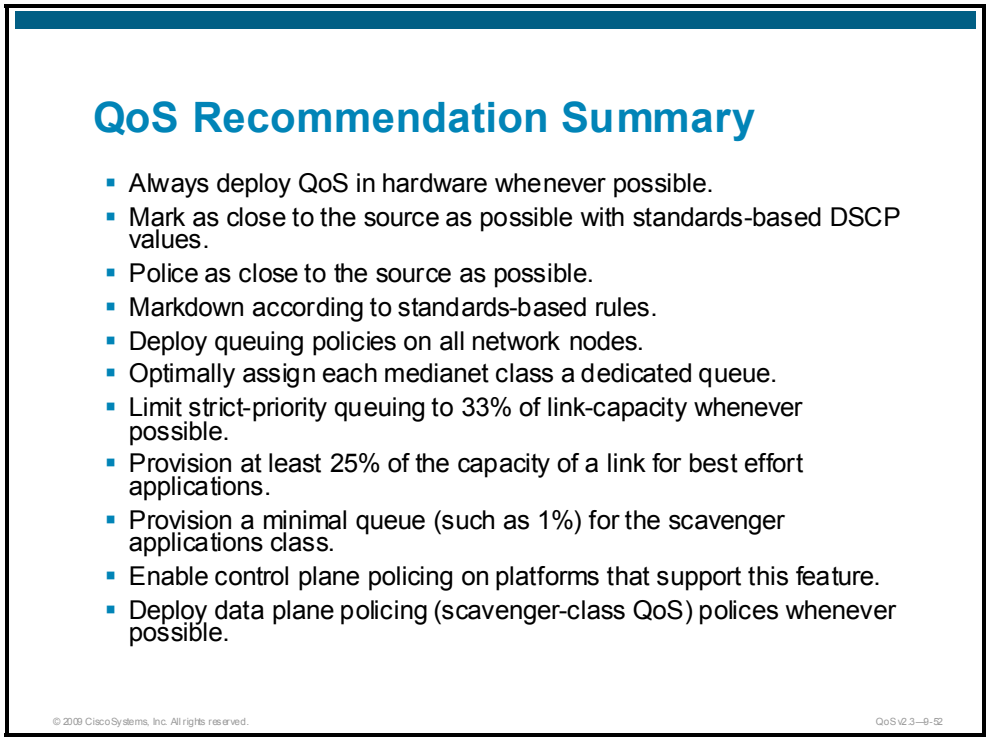
When these media services are made available within the network infrastructure, endpoints can be multi-purpose and rely upon these common media services to join and leave sessions for multiple media applications. Common functions such as transcoding and conferencing different media codecs within the same session can be deployed and leveraged by multiple applications, instead of being duplicated for each new media application.



The 12 classes of applications within this enterprise medianet QoS model, which have unique service level requirements and thus require explicit QoS PHBs, are outlined as follows:

- VoIP telephony
- Broadcast video
- Real-time interactive
- Multimedia conferencing
- Network control
- Signaling
- Operation, Administration, and Management (OAM)
- Transactional data and low-latency data
- Bulk data and high-throughput data
- Best effort
- Scavenger and low-priority data

Note For further information on the enterprise medianet QoS model, refer to the Enterprise Medianet Quality of Service Design 4.0 Overview on Cisco.com.



The following design principles can help simplify strategic QoS deployments:

Hardware versus Software QoS

A fundamental QoS design principle is to always enable QoS policies in hardware rather than software whenever a choice exists. Cisco Catalyst switches perform QoS in dedicated hardware ASICS on Ethernet-based ports and therefore do not tax their main CPUs to administer QoS policies. This allows complex policies to be applied at line rates at even gigabit or ten-gigabit speeds.

Classification and Marking Best Practices

When classifying and marking traffic, a recommended design principle is to classify and mark applications as close to their sources as technically and administratively feasible. This principle promotes end-to-end differentiated services and PHBs.

In general, it is not recommended to trust markings that can be set by users on their PCs or other similar devices, because users can easily abuse provisioned QoS policies if permitted to mark their own traffic. On the other hand, if enterprise controls are in place that centrally administer PC QoS markings, then it may be possible and advantageous to trust these.

Following this rule, it is further recommended to use DSCP markings whenever possible, because these are end-to-end, more granular, and more extensible than Layer 2 markings.

Policing and Markdown Best Practices

It is recommended to police traffic flows as close to their sources as possible. This principle applies also to legitimate flows, because worm-generated traffic can masquerade under legitimate, well-known TCP and UDP ports and cause extreme amounts of traffic to be poured onto the network infrastructure. Such excesses should be monitored at the source and marked down appropriately. Whenever supported, markdown should be done according to standards-based rules, such as RFC 2597 (AF PHB).

Queuing and Dropping Best Practices

Critical media applications require service guarantees regardless of network conditions. The only way to provide service guarantees is to enable queuing at any node that has the potential for congestion. This principle applies not only to campus-to-WAN and VPN edges, where speed mismatches are most pronounced, but also to campus interswitch links, where oversubscription ratios create the potential for congestion. There is simply no other way to guarantee service levels than by enabling queuing wherever a speed mismatch exists.

Additionally, because each medianet application class has unique service level requirements, each should optimally be assigned a dedicated queue. However, on platforms bounded by a limited number of hardware or service provider queues, no fewer than four queues would be required to support medianet QoS policies, specifically:

- Real-time queue (to support a RFC 3246 EF PHB service)
- Guaranteed-bandwidth queue (to support RFC 2597 AF PHB services)
- Default queue (to support a RFC 2474 DF service)
- Bandwidth-constrained queue (to support a RFC 3662 Scavenger service)

Strict-Priority Queuing Recommendations: The 33% LLQ Rule

If the majority of bandwidth is provisioned with strict priority queuing (which is effectively a FIFO queue), the overall effect is a dampening of QoS functionality, both for latency and jitter sensitive real-time applications (contending with each other within the FIFO priority queue) and also for non-real-time applications (as these may periodically receive wild bandwidth allocation fluctuations, depending on the instantaneous amount of traffic being serviced by the priority queue). Any traffic assigned to a strict-priority queue should be governed by an admission-control mechanism.

Best-Effort Queuing Recommendation

Because most enterprises have several thousand applications running over their networks, adequate bandwidth must be provisioned for this class as a whole in order to handle the sheer number and volume of applications that default to it. Therefore, it is recommended to reserve at least 25 percent of link bandwidth for the default best-effort class.

Scavenger Class Queuing Recommendations

Whenever scavenger queuing class is enabled, it should be assigned a minimal amount of bandwidth, such as 1 percent (or whatever the minimal bandwidth allocation that the platform supports). On some platforms, queuing distinctions between bulk data and scavenger traffic flows cannot be made. In such cases, the scavenger/bulk queue can be assigned a moderate amount of bandwidth, such as 5 percent.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- An SLA stipulates the delivery and pricing of numerous service levels. SLAs cover an assortment of data services such as Frame Relay, leased lines, Internet access, web hosting, etc.
- There are QoS requirements for different traffic types. Both the enterprise and the service provider must implement the proper QoS mechanisms to provide end-to-end QoS.
- These QoS design principles are important when deploying campus QoS policies:
  - Classify and mark applications as close to their sources as technically and administratively feasible.
  - Police unwanted traffic flows as close to their sources as possible.
  - Always perform QoS in hardware rather than software when a choice exists.
- You should enforce trust boundaries as close to the endpoints as technically and administratively possible; the definition of the trust boundary depends on the capabilities of the endpoints that are being connected to the access edge of the LAN.

Summary (Cont.)

- On WAN edges, queuing policies, shaping policies, selective dropping policies, and link-efficiency policies are typically required.
- For the most part, QoS design recommendations for WAN aggregators also apply to branch routers located at the far end of the WAN links; however, there are a few unique considerations for branch router QoS design.
- The service provider IP core is used to provide high-speed packet transport. Therefore, all markings, policing, and shaping should be performed only at the PE router on the PE-to-CE link and not at the core. Using the DiffServ model, only the edge requires complex QoS policy. At the core, only queuing and WRED are required. The operation of queuing and WRED will be based on the markings done at the edge (PE).

Summary (Cont.)

- To achieve end-to-end service levels over MPLS VPNs, enterprise and service-provider QoS designs must be consistent and complimentary.
- General guidelines for enterprise QoS implementations include the following:
  - Deploy QoS in hardware whenever possible.
  - Mark as close to the source as possible with standards-based DSCP values.
  - Police as close to the source as possible.
  - Markdown according to standards-based rules.
  - Deploy queuing policies on all network nodes.
  - Assign each medianet class a dedicated queue.
  - Limit strict-priority queuing to 33% of link-capacity.
  - Provision at least 25% of the capacity of a link for best effort applications.
  - Provision a minimal queue for the Scavenger applications class.
  - Use control plane policing and data plane policing.

Lesson 3

Providing QoS for Security

Overview

Denial of service (DoS) and worm attacks are increasing in frequency, complexity, and scope of damage. Quality of service (QoS) tools and strategic designs can mitigate the effects of worms and keep critical applications available during DoS attacks. One such strategy, referred to as scavenger-class QoS, uses a tactical approach to provide anomaly detection and reaction to DoS resulting from worm attack-generated traffic. This lesson describes how QoS tools can mitigate DoS attacks through the use of control plane, data plane, and Network-Based Application Recognition (NBAR) known-worm policing. This lesson does not provide the best practices in step-by-step configuration, but rather provides the reasons for deploying the policers.

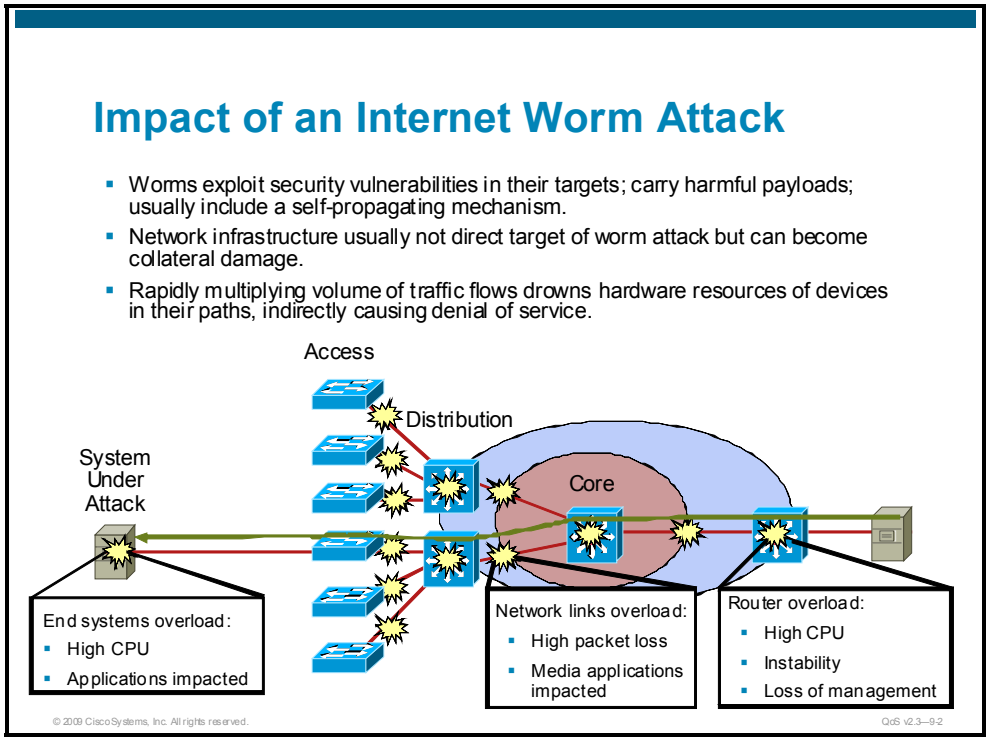
Objective

Upon completing this lesson, you will be able to describe steps recommended by Cisco to mitigate DoS attacks and worm attacks using QoS tools. This ability includes being able to meet these objectives:

- Describe tools and tactics used to increase security
- Explain how Control Plane Policing can be used to mitigate DoS attacks
- Explain how scavenger-class QoS can be used to mitigate DoS and Internet worm attacks
- Explain how NBAR can be used to mitigate Internet worms

QoS Tools and Tactics for Security

This topic describes the QoS tools that can be used to defend networks.



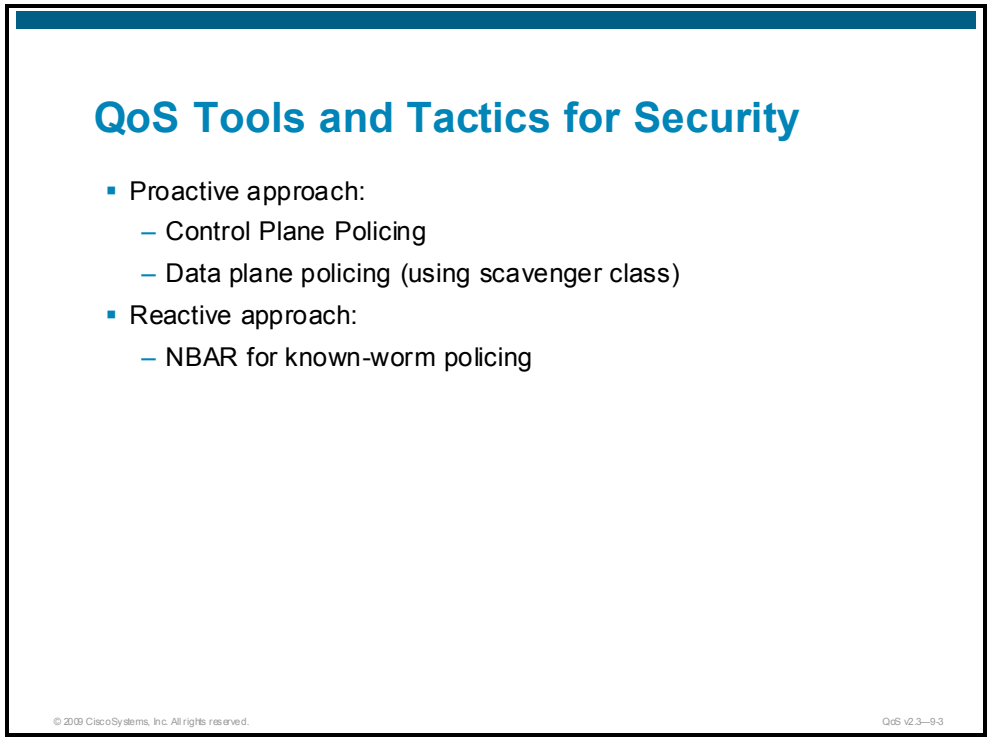
While the primary objective of most QoS deployments is to provision preferential, and sometimes deferential, service to various application classes, QoS policies can also provide a additional layer of security to the network infrastructure, especially in the case of mitigating denial of service (DoS) and worm attacks.

There are two main classes of DoS attacks:

- **Spoofing attacks:** The attacker pretends to provide a legitimate service, but provides false information or no information to the requester.
- **Slamming attacks:** The attacker exponentially generates and propagates traffic until service resources such as servers and network infrastructure are overwhelmed.

Spoofing attacks are best addressed by authentication and encryption technologies. Slamming (also known as "flooding") attacks, on the other hand, can be effectively mitigated through QoS technologies.

In contrast, worms exploit security vulnerabilities in their targets and carry harmful payloads that usually include a self-propagating mechanism. Network infrastructure usually is not the direct target of a worm attack but can become collateral damage as worms exponentially self-propagate. The rapidly multiplying volume of traffic flows eventually drowns the hardware resources of routers and switches in their paths, indirectly causing denial of service to legitimate traffic flows, as shown in the figure.



A reactive approach to mitigating such attacks is to reverse-engineer the worm and set up intrusion detection mechanisms, ACLs, or NBAR policies to limit its propagation. However, the increased sophistication and complexity of worms make them harder and harder to separate from legitimate traffic flows. This exacerbates the finite time lag between when a worm begins to propagate and when the following events can take place:

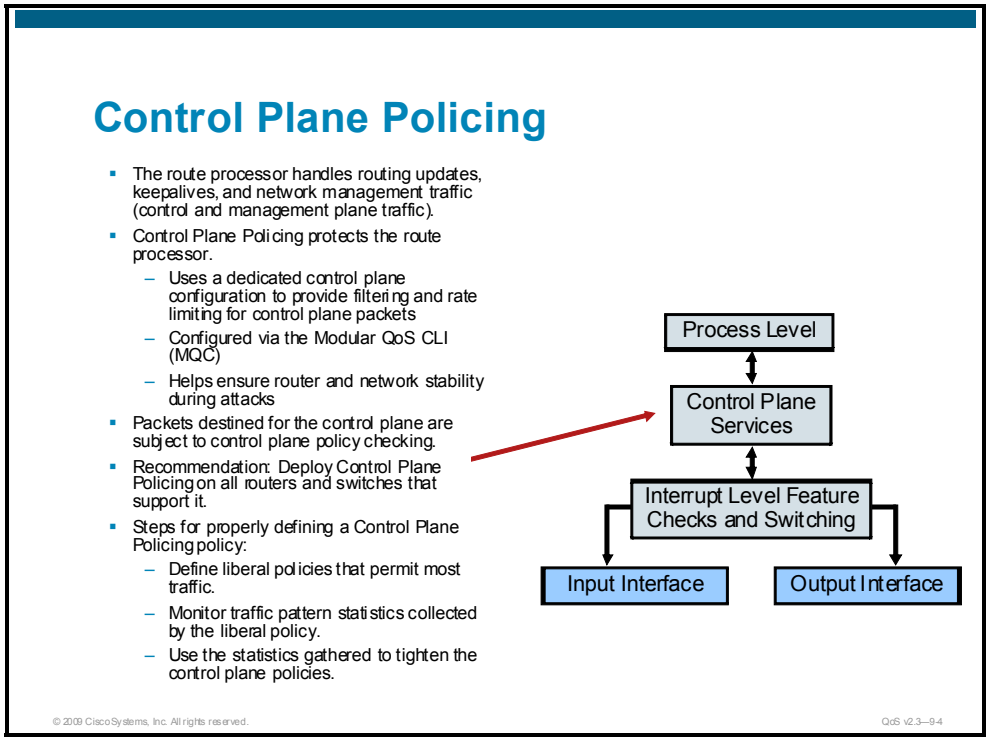
- Sufficient analysis has been performed to understand how the worm operates and what its network characteristics are.
- An appropriate patch, plug, or ACL is disseminated to network devices that may be in the path of worm; this task may be hampered by the attack itself, as network devices may become unreachable for administration during the attacks.

These time lags may not seem long in absolute terms, such as in minutes, but the relative window of opportunity for damage is huge. For example, in 2003, the number of hosts infected with the Slammer worm (a Sapphire worm variant) doubled every 8.5 seconds on average, infecting over 75,000 hosts in just 11 minutes and performing scans of 55 million more hosts within the same time period.

A proactive approach to mitigating DoS and worm attacks within enterprise networks is to have control plane policing and data plane policing policies in place within the infrastructure which immediately respond to out-of-profile network behavior indicative of DoS or worm attacks. Control Plane Policing serves to protect the CPU of network devices such as switches and routers from becoming bogged down with interruption-handling and thus not having enough cycles to forward traffic. Data plane policing (also referred to as Scavenger-class QoS) protects link bandwidth from being consumed by forwarding DoS worm traffic.

Control Plane Policing

This topic explains how Control Plane Policing can be used to increase network security.



A router or switch can be logically divided into four functional components or planes:

- Data plane
- Management plane
- Control plane
- Services plane

The vast majority of traffic travels through the router via the data plane. However the route processor must handle certain packets, such as routing updates, keepalives, and network management. This is often referred to as control and management plane traffic.

Because the route processor is critical to network operations, any service disruption to the route processor or the control and management planes can result in business-impacting network outages. A DoS attack targeting the route processor, which can be perpetrated either inadvertently or maliciously, typically involves high rates of traffic that result in excessive CPU utilization on the route processor itself. This type of attack, which can be devastating to network stability and availability, may display the following symptoms:

- Route processor CPU utilization is high. (near 100 percent)
- Line protocol keepalives and routing protocol updates are lost, leading to route flaps and major network transitions.
- Interactive sessions via the Command Line Interface (CLI) are slow or completely unresponsive due to high CPU utilization.

- Route processor resource exhaustion—resources such as memory and buffers are unavailable for legitimate IP data packets
- Packet queue backup, which leads to indiscriminate drops (or drops due to lack of buffer resources) of other incoming packets

Control Plane Policing addresses the need to protect the control and management planes, ensuring routing stability, availability, and packet delivery. It uses a dedicated control plane configuration via the Modular QoS CLI (MQC) to provide filtering and rate limiting capabilities for control plane packets.

The figure illustrates the flow of packets from various interfaces. Packets destined to the control plane are subject to control plane policy checking, as depicted by the control plane services block.

By protecting the route processor, Control Plane Policing helps ensure router and network stability during an attack. For this reason, a best-practice recommendation is to deploy Control Plane Policing as a key protection mechanism on all routers and switches that support it.

To successfully deploy Control Plane Policing, the existing control and management plane access requirements must be understood. While it can be difficult to determine the exact traffic profile required to build the filtering lists, the following summarizes the recommended steps necessary to properly define a Control Plane Policing policy:

1. Start the deployment by defining liberal policies that permit most traffic.
2. Monitor traffic pattern statistics collected by the liberal policy.
3. Use the statistics gathered in the previous step to tighten the control plane policies.

Control-Plane Policing Example

```
access-list 140 deny tcp host 10.1.1.1 any eq telnet
access-list 140 deny tcp host 10.1.1.2 any eq telnet
access-list 140 permit tcp any any eq telnet
!
class-map telnet-class
  match access-group 140
!
policy-map control-plane-in
  class telnet-class
    police 80000 conform transmit exceed drop
!
control-plane
  service-policy input control-plane-in
```

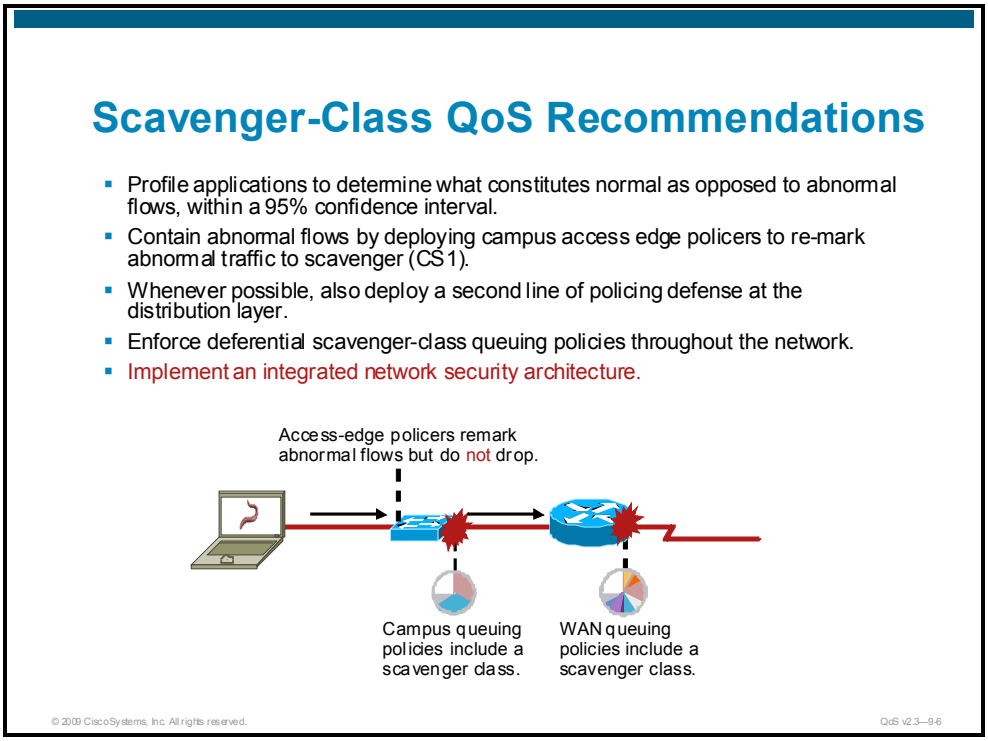

The example in the figure shows how to apply a QoS policy for aggregate control plane services to Telnet traffic that is received on the control plane. Trusted hosts with source addresses 10.1.1.1 and 10.1.1.2 are allowed to forward Telnet packets to the control plane without constraint; all other Telnet packets are policed at the specified rate.

In this example, the **access-list** commands that make up ACL 140 identify all Telnet traffic except Telnet traffic from trusted hosts 10.1.1.1 and 10.1.1.2. The **class-map** command defines a class-map called TELNET-CLASS, which is configured to match ACL 140. ACL 140 traffic is then associated with the policing action in policy-map CONTROL-PLANE-IN. The **control-plane** command enters control-plane configuration mode where the **service-policy** command is used to apply the CONTROL-PLANE-IN policy for aggregate control plane services to all packets that are entering the control plane from all line cards in the router.

Note	The use of the **deny** rule in access lists used in MQC is somewhat different from regular interface ACLs. Packets that match a **deny** rule are excluded from that class. This is in contrast to packets matching a **permit** rule, which are included in that class.

Data Plane Policing

This topic describes how to use policing and the scavenger class to protect the data plane.



The logic applied to protecting the control plane can also be applied to the data plane. Data plane policing (scavenger-class QoS) has two components, as illustrated in the figure:

- Campus access-edge policers that meter traffic flows from endpoint devices and remark abnormal flows to CS1, the scavenger marking value.
- Queuing policies on all nodes that include a deferential service class for scavenger traffic.

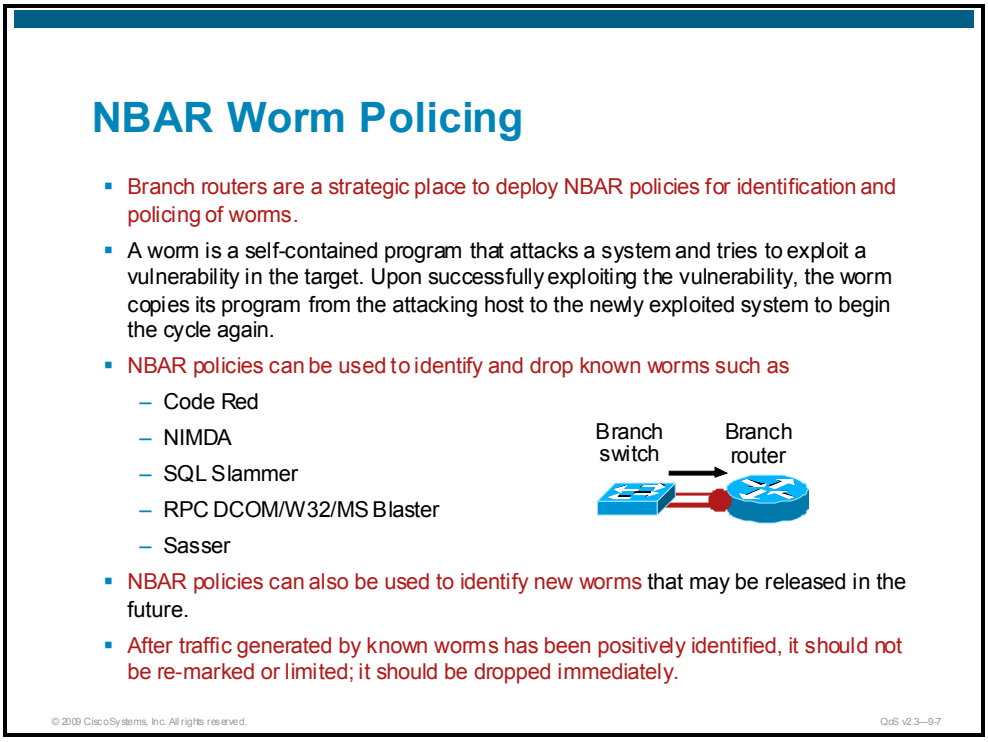
To implement data plane policing, you must first profile applications to determine what constitutes normal as opposed to abnormal flows, within a 95 percent confidence interval. Thresholds demarking normal and abnormal flows vary from enterprise to enterprise and from application to application. Beware of over-scrutinizing traffic behavior because this could exhaust time and resources and could easily change daily. Remember, legitimate traffic flows that temporarily exceed thresholds are not penalized by the data plane policing strategy. Only sustained, abnormal streams generated simultaneously by multiple hosts (highly indicative of DoS or worm attacks) are subject to aggressive dropping only after legitimate traffic has been serviced.

To contain such abnormal flows, deploy campus access edge policers to remark abnormal traffic to scavenger (CS1). Additionally, whenever possible, deploy a second line of policing defense at the distribution layer. To complement these remarking policies, enforce deferential scavenger class queuing policies throughout the network.

It is important to recognize the distinction between mitigating an attack and preventing it entirely. Control plane policing and data plane policing policies do not guarantee that DoS or worm attacks will never happen, but serve only to reduce the risk and the impact that such attacks could have on the network infrastructure. Therefore, it is vital to overlay a comprehensive security strategy on the QoS-enabled network infrastructure.

NBAR Worm Policing

This topic describes the use of NBAR to suppress worm attacks.



Branch routers are a strategic place to deploy NBAR policies for worm identification and policing. NBAR policies can be used to identify and drop Code Red, NIMDA, SQL Slammer, RPC DCOM/W32/MS Blaster, Sasser, and other known worms. NBAR policies can also be used to identify new worms that may be released in the future.

Typically, worms are self-contained programs that attack a system and try to exploit a vulnerability in the target. Upon successfully exploiting the vulnerability, the worm copies its program from the attacking host to the newly exploited system to begin the cycle again.

Note	A virus, which is slightly different from a worm, requires a vector to carry the virus code from one system to another. The vector can be either a word-processing document, an email message, or an executable program. The main element that distinguishes a worm from a virus is that a computer virus requires human intervention to facilitate its spreading, whereas worms (once released) propagate without requiring additional human intervention.

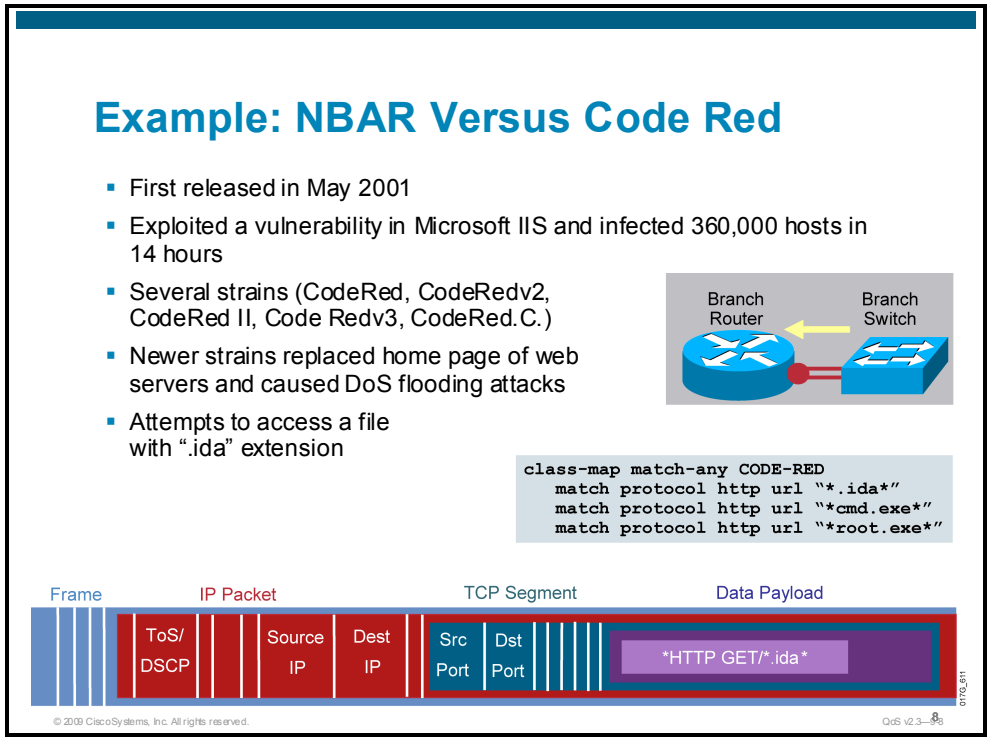
Worms are comprised of three primary components:

- **The enabling exploit code:** The enabling exploit code is used to exploit a vulnerability on a system. Exploitation of this vulnerability provides access to the system and the capability to execute commands on the target system.
- **A propagation mechanism:** When access has been obtained through the enabling exploit, the propagation mechanism is used to replicate the worm to the new target. The method used to replicate the worm can be achieved through the use of the Trivial File Transfer Protocol (TFTP), FTP, or another communication method. When the worm code is brought to the new host, the cycle of infection can be started again.

- **A payload:** Some worms also contain payloads, which might include additional code to further exploit the host, modify data on the host, or change a web page. A payload is not a required component, and, in many cases, the worm's enabling exploit code itself can be considered the payload.

Some worms use unique TCP and UDP ports to propagate. These types of worms are fairly simple to block using access lists (when the ports are known). Such ACLs can be configured on the branch switch (whenever supported) or on the LAN edge of the branch router.

Other worms hijack legitimate TCP and UDP ports to carry their harmful payloads. For these latter types of worms, NBAR can be used at the branch LAN edge to perform deep-packet analysis and drop any packets that are carrying the payloads of known worms. Code Red, NIMDA, SQL Slammer, RPC DCOM/W32/MS Blaster, and Sasser are examples of known worms.



NBAR Versus Code Red

First released in July 2001, Code Red targeted Microsoft Internet Information Server (IIS) using a vulnerability in the IIS Indexing Service. Although the first variant of this worm did little damage because of a flaw in the random number-generator code used to generate addresses of hosts to exploit, a second variant appeared with the flaw fixed.

This worm, CodeRedv2, spread quickly and became the most widespread and damaging worm to hit the Internet since the Morris worm. The success of CodeRedv2 as a worm relied on the fact that the worm exploited the vulnerability in the IIS Indexing Service only as a means of gaining access to the host. This, coupled with the wide deployment of IIS as well as the large number of unpatched IIS web servers, contributed to the quick and wide-ranging spread of the worm. An estimated 360,000 hosts were infected within a period of 14 hours.

CodeRedv2 temporarily replaced the home page of the web servers that it struck with a new page. Additionally, the code of the worm indicated that it was programmed to begin a packet-flooding DoS attack against a hard-coded IP address. (At the time, this was the IP address of the White House web server at http://www.whitehouse.gov.)

The initial infection attempt sends a large HTTP GET request to the target IIS server. In both the original Code Red and CodeRedv2, the GET request looks for a file named default.ida. However, this filename changes in newer variants, such as CodeRedv3/CodeRed.C. Although the filename has changed, the .ida suffix remains the same.

Code Red variants can include payloads that execute cmd.exe or root.exe functions to program scripts within the IIS scripts directory, thus providing a ready-made back door to the server for any attacker to use. Therefore, to combat Code Red, NBAR policies can be configured to check the payload of HTTP packets for .ida, cmd.exe, and root.exe, as shown in the figure.

NBAR Versus SQL Slammer

In January 2003, a new worm infected the Internet at such a high rate that it was categorized as a flash worm. This worm, termed SQL Slammer, once again targeted Microsoft Windows servers; specifically, this worm targeted servers running Microsoft Structured Query Language (SQL) Server software. The vulnerability exploited by SQL Slammer had been published in July 2002, and a patch from Microsoft was available at that time as well. Even though this patch was available for almost six months, SQL Slammer spread with incredibly high efficiency.

SQL slammer is a 376-byte User Datagram Protocol (UDP)-based worm that infects Microsoft SQL servers through UDP port 1434. Because of its small size, the SQL Slammer worm is contained in a single packet. The fast scanning rate of SQL Slammer is achieved not only because of this small size, but also because the worm is UDP based. The worm does not have to complete a handshake to connect with a target system.

SQL Slammer reached its full scanning rate of 55 million scans per second within 3 minutes of the start of the infection and infected the majority of vulnerable hosts on the Internet within 10 minutes of the start of the infection, with an estimated 300,000 infected hosts overall. A major consequence of such a fast scanning rate was that edge networks were overwhelmed by the amount of traffic generated by the worm. The doubling rate of SQL Slammer was approximately 8.5 seconds. In contrast, the doubling rate of CodeRedv2 was about 37 minutes.

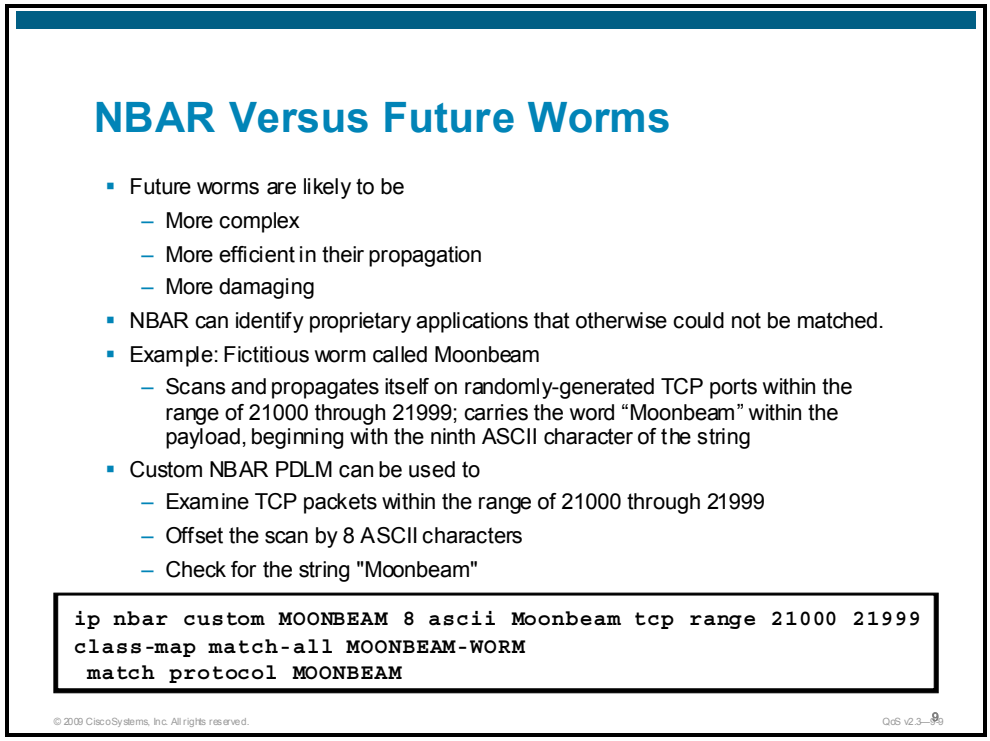
SQL Slammer does not carry an additional harmful payload (beyond its enabling exploit code), and its primary purpose is to cause DoS through exponential self propagation.

NBAR can be used to detect the SQL Slammer worm by mapping a custom PDLM to UDP port 1434 and matching on the packet length (376-byte worm + 8 bytes of UDP header + 20 bytes of IP header = 404 bytes). This is shown in the following example:

```
ip nbar port-map custom-02 udp 1434
class-map match-all SQL-SLAMMER
   match protocol custom-02
   match packet length min 404 max 404
```

You can use the **show policy** command and the **show ip nbar port-map** command to verify your configuration.

Note For further information on NIMDA, Sasser, and RPC DCOM/W32/MS Blaster, as well as details on how to configure NBAR policies that identify and drop these known worms, refer to the “Branch Router QoS Design” section of the *Enterprise QoS Solution Reference Network Design Guide* on Cisco.com.



There is every reason to believe that new worms will be released in the future. These worms will be not only more complex, but also more efficient in their propagation, and thus more damaging in their scope.

An NBAR feature that was introduced in Cisco IOS Software Release 12.3[4]T enables you to extend the capability of NBAR to classify and monitor additional static port applications or to allow NBAR to classify unsupported static port traffic. Specifically, it enables you to define the strings that you want to search for in the application payload in order to identify a given application. This applies to any application, not just HTTP URLs.

This functionality can be used to identify proprietary applications that otherwise could not be matched. However, it also can be very useful in plugging holes that future worms might open.

The figure shows an example that uses a fictitious worm called Moonbeam. Moonbeam scans and propagates itself on randomly generated TCP ports within the range of 21000 through 21999. Furthermore, the worm carries the word Moonbeam within the payload, beginning with the ninth ASCII character of the string. The payload of Moonbeam might look like the following:

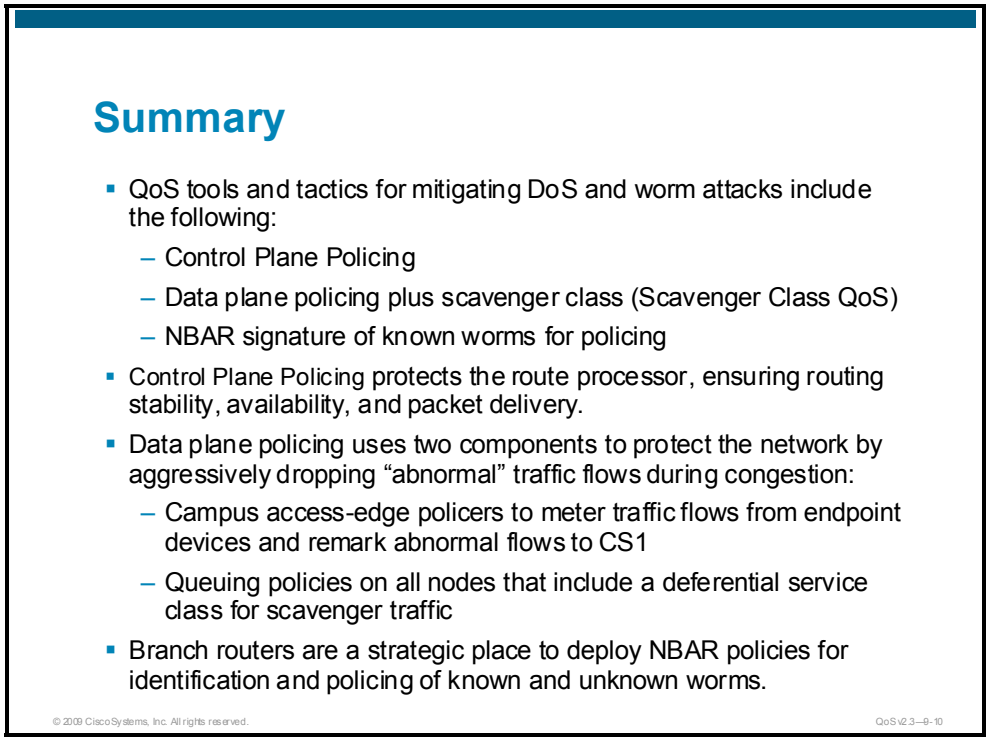
```
%u[65&%]Moonbeam\x01\x01\x01\x01.*[.][Dd][Ll][Ll]u9090%u6858%ucbd3%
u7801%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u9090%
u8190%u00c3%u0003%u8b00%u531b%u53ff%u0078%u0000%u00=a 403
```

Moonbeam could be identified by a custom NBAR PDLM that examines TCP packets within the range of 21000 through 21999, offsets the scan by 8 ASCII characters, and checks for the string "Moonbeam" (case sensitive), as shown in the figure.

You can use the **ip nbar custom** command in global configuration mode to create the custom NBAR PDLM. You can then use the **show policy** and show **ip nbar port-map** commands to verify your configuration. The **ip nbar port-map** command displays the current protocol-to-port mappings in use by NBAR.

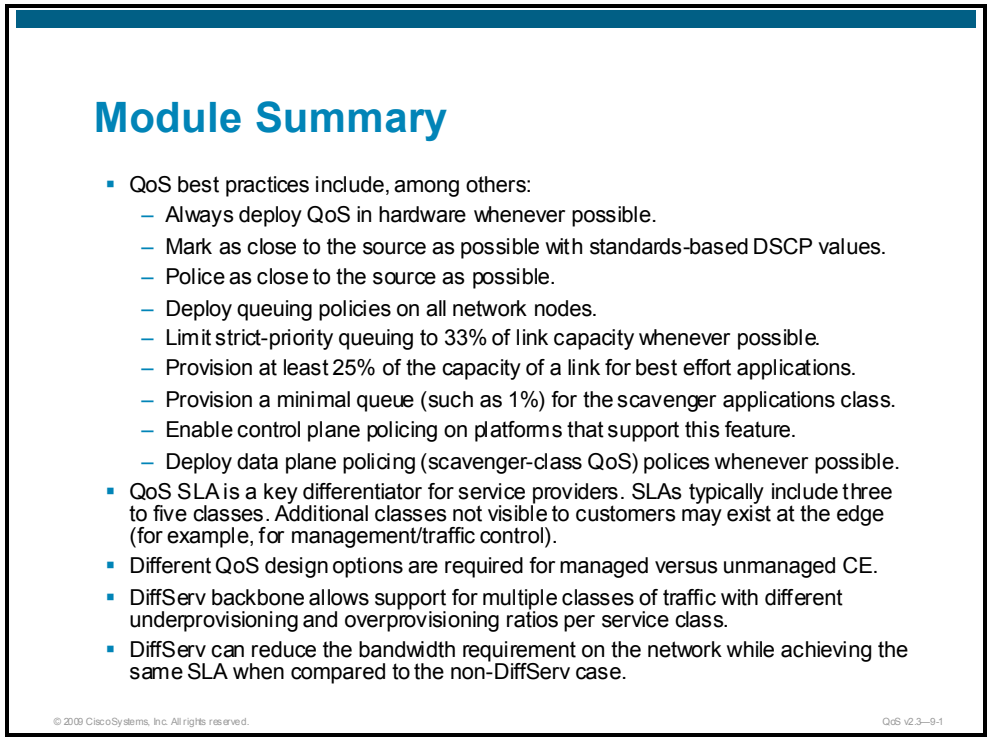
Summary

This topic summarizes the key points discussed in this lesson.



Module Summary

This topic summarizes the key points discussed in this module.



When an enterprise network is connected by a service provider that provides Layer 3 IP services, both the enterprise and the service provider must implement the proper quality of service (QoS) mechanisms. This is necessary to satisfy the end-to-end QoS requirements of the different applications running at the enterprise. A key differentiator for service providers when offering Layer 3 IP services is to offer service level agreements (SLAs) for each of the traffic classes supported by the service provider. SLAs provide the needed assurance to the enterprise customers that the service provider network can meet the QoS requirements of the enterprise applications traversing the service provider network.

This module discusses some of the key best practices for deploying QoS at the enterprise and at the service provider network. The first lesson of the module includes a baseline traffic classification recommendation detailing the per-hop behavior (PHB), differentiated services code point (DSCP), IP precedence, and Layer 2 class of service (CoS) values for some of the common enterprise traffic types. The traffic classification recommendation serves as a guide for overall system design and QoS feature implementation.

The second lesson of the module provides a case study illustrating typical QoS implementations at the enterprise campus LAN, at the WAN edge (customer edge [CE] and provider edge [PE] routers), and at the service provider core.

Within the enterprise campus LAN, a hierarchical design is required along with proper buffer management at the LAN switches. Classification and marking should be performed as close to the source as possible. Class-based policing can be implemented to rate-limit certain non-business-related applications (like peer-to-peer file-sharing applications, Napster, and so on).

More complex QoS configurations are required at the WAN edge. The QoS configuration required at the WAN edge CE and PE routers is different depending on whether or not the CE is managed by the service provider or by the enterprise customer. Low-latency queuing (LLQ) and class-based weighted fair queuing (CBWFQ) are required for congestion management. Traffic shaping is required on Frame Relay permanent virtual circuits (PVCs). Link efficiency mechanisms like link fragmentation and interleaving (LFI) and compressed Real-Time Transport Protocol (cRTP) are required to improve the WAN link efficiency. Class-based marking is also required if the enterprise customer traffic classes need to be remapped into the service provider traffic classes.

At the service provider high-speed core, typically only weighted random early detection (WRED), LLQ, or modified deficit round robin (MDRR) are required, based on the previous packet QoS marking. In the service provider core (backbone), two of the design options include implementing a best-effort backbone with overprovisioning, or implementing a DiffServ backbone with a different provisioning ratio for each traffic class. By deploying a DiffServ backbone, service providers can significantly reduce their cost while providing the required SLAs for their customers.

References

For additional information on Enterprise QoS Design, refer to *Cisco Enterprise Quality of Service Solution Reference Network Design Guide* at this URL: http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND/QoSIntro.html

Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) Which three are QoS requirements for voice? (Choose three.) (Source: Understanding Traffic Classification Best Practices)

A) latency ≤ 150 ms
B) jitter < 30 ms
C) loss ≤ 1 percent
D) bandwidth = 80 kbps per call minimum for all codecs
E) retransmit interval ≤ 10 ms

Q2) Which two statements are general guidelines for implementing campus QoS? (Source: Deploying End-to-End QoS)

A) Police unwanted traffic flows as close to their destinations as possible.
B) Classification or marking should be done at the high-speed core layer.
C) Always use NBAR to classify traffic because all Cisco Catalyst switches support NBAR.
D) Classify and mark applications as close to their sources as technically and administratively feasible.
E) QoS in the campus is generally not a buffer management issue as much as it is a bandwidth management issue. Therefore, link efficient mechanisms should be implemented between the access and distribution layer links.

Q3) Based on the following configuration, which two statements are true? (Choose two.) (Source: Deploying End-to-End QoS)

```
class-map match-all PREMIUM
      match ip access-group 101
    class-map match-all BUSINESS
      match ip access-group 102
    !
    policy-map OUT-POLICY
      class PREMIUM
       priority percent 25
       set ip dscp ef
      class BUSINESS
       bandwidth percent remaining 75
       set ip dscp af31
       random-detect dscp-based
      class class-default
       bandwidth percent remaining 25
       set ip dscp 0
       random-detect dscp-based
```

A) The premium traffic class has a maximum bandwidth guarantee equal to 25 percent of the available interface bandwidth.
B) The business traffic class has a maximum bandwidth guarantee equal to 75 percent of the interface bandwidth.
C) The default traffic class has a maximum bandwidth guarantee equal to 25 percent of the interface bandwidth.
D) Tail drop is used on the business traffic class.
E) WRED is used on the premium traffic class.
F) WRED is used on the default traffic class.

Q4) The QoS requirements on the CE and PE router differ depending on which factor? (Source: Deploying End-to-End QoS)

A) whether or not the PE router is managed by the service provider
B) whether or not the CE router is managed by the service provider
C) whether or not the service provider is using an MPLS core
D) the number of traffic classes supported by the service provider
E) the SLAs offered by the service provider

Module Self-Check Answer Key

Q1) A, B, C

Q2) D

Q3) A, F

Q4) B