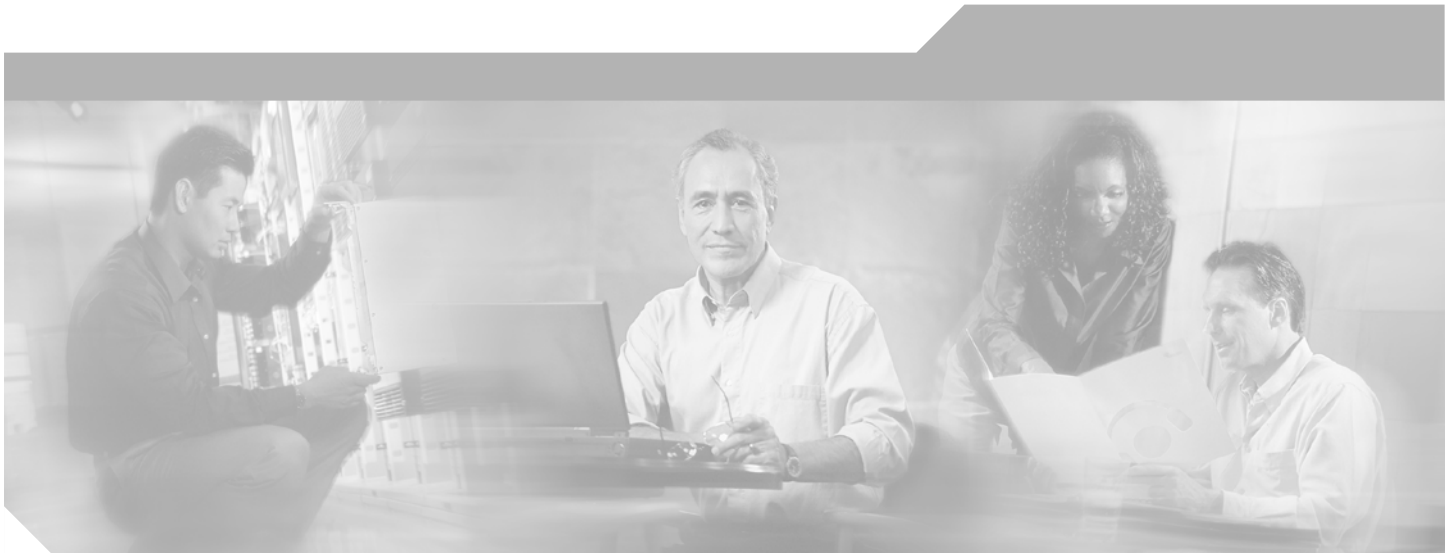
Cisco Systems



Enterprise QoS Solution Reference Network Design Guide

Version 3.3
November 2005

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

CONTENTS

Preface

This document provides design considerations and guidelines for implementing Cisco Quality of Service within an enterprise environment.

This document is the second major update to the design guidelines and information presented in the *Cisco AVVID Network Infrastructure Enterprise Quality of Service Design Solutions Reference Network Design* (August, 2002).

This document assumes that you are already familiar with Quality of Service terms and concepts. If you want to review any of those terms and concepts, refer to Cisco Quality of Service documentation at:

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

Alternatively, Quality of Service tools and concepts are presented in depth within the Cisco Press book *End-to-End Quality of Service Network Design* (ISBN: 1587051761).

Revision History

The following table lists the revision history for this document.

Revision Date	Comments
November 2005	Revision 3.3 with terminology change from "Business Ready" to "Enterprise."
October 2005	The following section has been added: • IPSec VPN QoS Design—Chapter 6
June 2005	The following sections are new or have been added • AutoQoS—VoIP (Campus)—Chapter 2 • AutoQoS—Enterprise (WAN)—Chapter 4 • Technical corrections and edits
April 2005	Initial Draft (QoS SRND 3.0)

Obtaining Documentation

Cisco provides several ways to obtain documentation, technical assistance, and other technical resources. These sections explain how to obtain technical information from Cisco Systems.

Cisco.com

You can access the most current Cisco documentation on the World Wide Web at this URL:

http://www.cisco.com/univercd/home/home.htm

You can access the Cisco website at this URL:

http://www.cisco.com

International Cisco web sites can be accessed from this URL:

http://www.cisco.com/public/countries_languages.shtml

Documentation CD-ROM

Cisco documentation and additional literature are available in a Cisco Documentation CD-ROM package, which may have shipped with your product. The Documentation CD-ROM is updated monthly and may be more current than printed documentation. The CD-ROM package is available as a single unit or through an annual subscription.

Registered Cisco.com users can order the Documentation CD-ROM (product number DOC-CONDOCCD=) through the online Subscription Store:

http://www.cisco.com/go/subscription

Ordering Documentation

You can find instructions for ordering documentation at this URL:

http://www.cisco.com/univercd/cc/td/doc/es_inpck/pdi.htm

You can order Cisco documentation in these ways:

- Registered Cisco.com users (Cisco direct customers) can order Cisco product documentation from the Networking Products MarketPlace:

  http://www.cisco.com/en/US/partner/ordering/index.shtml

- Registered Cisco.com users can order the Documentation CD-ROM (Customer Order Number DOC-CONDOCCD=) through the online Subscription Store:

  http://www.cisco.com/go/subscription

- Nonregistered Cisco.com users can order documentation through a local account representative by calling Cisco Systems Corporate Headquarters (California, U.S.A.) at 408 526-7208 or, elsewhere in North America, by calling 800 553-NETS (6387).

Documentation Feedback

You can submit comments electronically on Cisco.com. On the Cisco Documentation home page, click Feedback at the top of the page.

You can e-mail your comments to bug-doc@cisco.com.

You can submit your comments by mail by using the response card behind the front cover of your document or by writing to the following address:

Cisco Systems

Attn: Customer Document Ordering

170 West Tasman Drive

San Jose, CA 95134-9883

We appreciate your comments.

Obtaining Technical Assistance

Cisco provides Cisco.com, which includes the Cisco Technical Assistance Center (TAC) Website, as a starting point for all technical assistance. Customers and partners can obtain online documentation, troubleshooting tips, and sample configurations from the Cisco TAC website. Cisco.com registered users have complete access to the technical support resources on the Cisco TAC website, including TAC tools and utilities.

Cisco.com

Cisco.com offers a suite of interactive, networked services that let you access Cisco information, networking solutions, services, programs, and resources at any time, from anywhere in the world. Cisco.com provides a broad range of features and services to help you with these tasks:

- Streamline business processes and improve productivity
- Resolve technical issues with online support
- Download and test software packages
- Order Cisco learning materials and merchandise
- Register for online skill assessment, training, and certification programs

To obtain customized information and service, you can self-register on Cisco.com at this URL:

http://www.cisco.com

Technical Assistance Center

The Cisco TAC is available to all customers who need technical assistance with a Cisco product, technology, or solution. Two levels of support are available: the Cisco TAC website and the Cisco TAC Escalation Center. The avenue of support that you choose depends on the priority of the problem and the conditions stated in service contracts, when applicable.

We categorize Cisco TAC inquiries according to urgency:

- Priority level 4 (P4)—You need information or assistance concerning Cisco product capabilities, product installation, or basic product configuration.
- Priority level 3 (P3)—Your network performance is degraded. Network functionality is noticeably impaired, but most business operations continue.
- Priority level 2 (P2)—Your production network is severely degraded, affecting significant aspects of business operations. No workaround is available.
- Priority level 1 (P1)—Your production network is down, and a critical impact to business operations will occur if service is not restored quickly. No workaround is available.

Cisco TAC Website

You can use the Cisco TAC website to resolve P3 and P4 issues yourself, saving both cost and time. The site provides around-the-clock access to online tools, knowledge bases, and software. To access the Cisco TAC website, go to this URL:

http://www.cisco.com/tac

All customers, partners, and resellers who have a valid Cisco service contract have complete access to the technical support resources on the Cisco TAC website. Some services on the Cisco TAC website require a Cisco.com login ID and password. If you have a valid service contract but do not have a login ID or password, go to this URL to register:

http://tools.cisco.com/RPF/register/register.do

If you are a Cisco.com registered user, and you cannot resolve your technical issues by using the Cisco TAC website, you can open a case online at this URL:

http://www.cisco.com/en/US/support/index.html

If you have Internet access, we recommend that you open P3 and P4 cases through the Cisco TAC website so that you can describe the situation in your own words and attach any necessary files.

Cisco TAC Escalation Center

The Cisco TAC Escalation Center addresses priority level 1 or priority level 2 issues. These classifications are assigned when severe network degradation significantly impacts business operations.

When you contact the TAC Escalation Center with a P1 or P2 problem, a Cisco TAC engineer automatically opens a case.

To obtain a directory of toll-free Cisco TAC telephone numbers for your country, go to this URL:

http://www.cisco.com/warp/public/687/Directory/DirTAC.shtml

Before calling, please check with your network operations center to determine the level of Cisco support services to which your company is entitled: for example, SMARTnet, SMARTnet Onsite, or Network Supported Accounts (NSA). When you call the center, please have available your service agreement number and your product serial number.

Obtaining Additional Publications and Information

Information about Cisco products, technologies, and network solutions is available from various online and printed sources.

- The Cisco Product Catalog describes the networking products offered by Cisco Systems as well as ordering and customer support services. Access the Cisco Product Catalog at this URL:

  http://www.cisco.com/en/US/products/products_catalog_links_launch.html

- Cisco Press publishes a wide range of networking publications. Cisco suggests these titles for new and experienced users: Internetworking Terms and Acronyms Dictionary, Internetworking Technology Handbook, Internetworking Troubleshooting Guide, and the Internetworking Design Guide. For current Cisco Press titles and other information, go to Cisco Press online at this URL:

  http://www.ciscopress.com

- Packet magazine is the Cisco monthly periodical that provides industry professionals with the latest information about the field of networking. You can access Packet magazine at this URL:

  http://www.cisco.com/en/US/about/ac123/ac114/about_cisco_packet_magazine.html

- iQ Magazine is the Cisco monthly periodical that provides business leaders and decision makers with the latest information about the networking industry. You can access iQ Magazine at this URL:

  http://business.cisco.com/prod/tree.taf%3fasset_id=44699&public_view=true&kbns=1.html

- Internet Protocol Journal is a quarterly journal published by Cisco Systems for engineering professionals involved in the design, development, and operation of public and private internets and intranets. You can access the Internet Protocol Journal at this URL:

  http://www.cisco.com/en/US/about/ac123/ac147/about_cisco_the_internet_protocol_journal.html

- Training-Cisco offers world-class networking training, with current offerings in network training listed at this URL:

  http://www.cisco.com/en/US/learning/le31/learning_recommended_training_list.html

Quality of Service Design Overview

This document provides an overview of Quality of Service (QoS) tools and design and includes high-level answers to the following questions:

- Why is Quality of Service Important for Enterprise Networks?
- What is Cisco's Quality of Service Toolset?
- How is QoS Optimally Deployed within an Enterprise?
- How can QoS Tools be used to Mitigate DoS/Worm Attacks?

QoS has already proven itself as the enabling technology for the convergence of voice, video and data networks. As business needs evolve, so do demands on QoS technologies. The need to protect voice, video and critical data via QoS mechanisms in an enterprise network has escalated over the past few years, primarily due to the increased frequency and sophistication of Denial of Service (DoS) and worm attacks. This document examines current QoS demands and requirements within the enterprise and presents strategic design recommendations to address these needs.

QoS Overview

This section answers the following questions:

- What is QoS?
- Why is QoS Important for Enterprise Networks?

What is QoS?

QoS is the measure of transmission quality and service availability of a network (or internetworks).

Service availability is a crucial foundation element of QoS. The network infrastructure must be designed to be highly available before you can successfully implement QoS. The target for High Availability is 99.999 % uptime, with only five minutes of downtime permitted per year. The transmission quality of the network is determined by the following factors:

- Loss—A relative measure of the number of packets that were not received compared to the total number of packets transmitted. Loss is typically a function of availability. If the network is Highly Available, then loss during periods of non-congestion would be essentially zero. During periods of congestion, however, QoS mechanisms can determine which packets are more suitable to be selectively dropped to alleviate the congestion.

- Delay—The finite amount of time it takes a packet to reach the receiving endpoint after being transmitted from the sending endpoint. In the case of voice, this is the amount of time it takes for a sound to travel from the speaker's mouth to a listener's ear.
- Delay variation (Jitter)—The difference in the end-to-end delay between packets. For example, if one packet requires 100 ms to traverse the network from the source endpoint to the destination endpoint and the following packet requires 125 ms to make the same trip, then the delay variation is 25 ms.

Each end station in a Voice over IP (VoIP) or Video over IP conversation uses a *jitter buffer* to smooth out changes in the arrival times of voice data packets. Although jitter buffers are dynamic and adaptive, they may not be able to compensate for instantaneous changes in arrival times of packets. This can lead to jitter buffer over-runs and under-runs, both of which result in an audible degradation of call quality.

Why is QoS Important for Enterprise Networks?

A communications network forms the backbone of any successful organization. These networks transport a multitude of applications, including realtime voice, high-quality video and delay-sensitive data. Networks must provide predictable, measurable, and sometimes guaranteed services by managing bandwidth, delay, jitter and loss parameters on a network.

QoS technologies refer to the set of tools and techniques to manage network resources and are considered the key enabling technology for network convergence. The objective of QoS technologies is to make voice, video and data convergence appear transparent to end users. QoS technologies allow different types of traffic to contend inequitably for network resources. Voice, video, and critical data applications may be granted priority or preferential services from network devices so that the quality of these strategic applications does not degrade to the point of being unusable. Therefore, QoS is a critical, intrinsic element for successful network convergence.

QoS tools are not only useful in protecting desirable traffic, but also in providing deferential services to undesirable traffic such as the exponential propagation of worms. You can use QoS to monitor flows and provide first and second order reactions to abnormal flows indicative of such attacks, as will be discussed in additional detail later in this document.

What is the Cisco QoS Toolset?

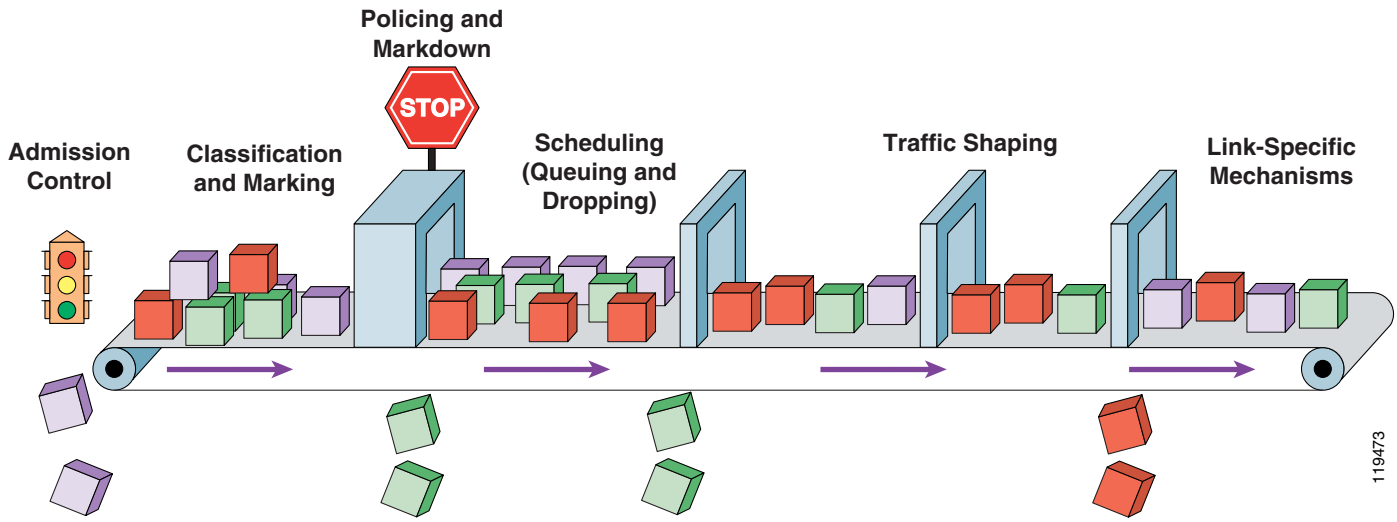
This section describes the main categories of the Cisco QoS toolset and includes the following topics:

- Classification and Marking tools
- Policing and Markdown tools
- Scheduling tools
- Link-specific tools
- AutoQoS tools
- Call Admission Control tools

Cisco provides a complete toolset of QoS features and solutions for addressing the diverse needs of voice, video and multiple classes of data applications. Cisco QoS technology lets complex networks control and predictably service a variety of networked applications and traffic types. You can effectively control bandwidth, delay, jitter, and packet loss with these mechanisms. By ensuring the desired results,

the QoS features lead to efficient, predictable services for business-critical applications. Using the rich Cisco QoS toolset, as shown in Figure 1-1, businesses can build networks that conform to the Differentiated Services (DiffServ) architecture, as defined in RFC 2475.

Figure 1-1 The Cisco QoS Toolset



Classification and Marking Tools

The first element to a QoS policy is to classify/identify the traffic that is to be treated differently. Following classification, marking tools can set an attribute of a frame or packet to a specific value. Such marking (or remarking) establishes a trust boundary that scheduling tools later depend on.

Classification and marking tools set this trust boundary by examining any of the following:

- Layer 2 parameters—802.1Q Class of Service (CoS) bits, Multiprotocol Label Switching Experimental Values (MPLS EXP)
- Layer 3 parameters—IP Precedence (IPP), Differentiated Services Code Points (DSCP), IP Explicit Congestion Notification (ECN), source/destination IP address
- Layer 4 parameters— L4 protocol (TCP/UDP), source/destination ports
- Layer 7 parameters— application signatures via Network Based Application Recognition (NBAR)

NBAR is a Cisco proprietary technology that identifies application layer protocols by matching them against a Protocol Description Language Module (PDLM), which is essentially an application signature. The NBAR deep-packet classification engine examines the data payload of stateless protocols against PDLMs. There are over 98 PDLMs embedded into Cisco IOS software 12.3 code. Additionally, Cisco IOS software 12.3(4)T introduces the ability to define custom PDLMs which examine user-defined strings within packet payloads. PDLMs can be added to the system without requiring an IOS upgrade because they are modular. NBAR is dependent on Cisco Express Forwarding (CEF) and performs deep-packet classification only on the first packet of a flow. The remainder of the packets belonging to the flow is then CEF-switched.

You can only apply policies to traffic after it has been positively classified. To avoid the need for repetitive and detailed classification at every node, packets can be marked according to their service levels. An analogy: imagine that each individual in the postal system would have to open up each letter to determine the respective priority required and service it accordingly. Obviously it would be better to

have the first mail-clerk stamp something on the outside of the envelope to indicate the priority level that would be applied during each phase of processing and delivery. Similarly, marking tools can be used to indicate respective priority levels by setting attributes in the frame/packet headers so that detailed classification does not have to be recursively performed at each hop. Within an enterprise, marking is done at either Layer 2 or Layer 3, using the following fields:

- 802.1Q/p Class of Service (CoS)—Ethernet frames can be marked at Layer 2 with their relative importance by setting the 802.1p User Priority bits of the 802.1Q header. Only three bits are available for 802.1p marking. Therefore, only 8 classes of service (0-7) can be marked on Layer 2 Ethernet frames.
- IP Type of Service (ToS) byte—Layer 2 media often changes as packets traverse from source to destination, so a more ubiquitous classification occurs at Layer 3. The second byte in an IPv4 packet is the ToS byte. The first three bits of the ToS byte are the IPP bits. These first three bits combined with the next three bits are known collectively as the DSCP bits.

  The IP Precedence bits, like 802.1p CoS bits, allow for only the following 8 values of marking (0–7):

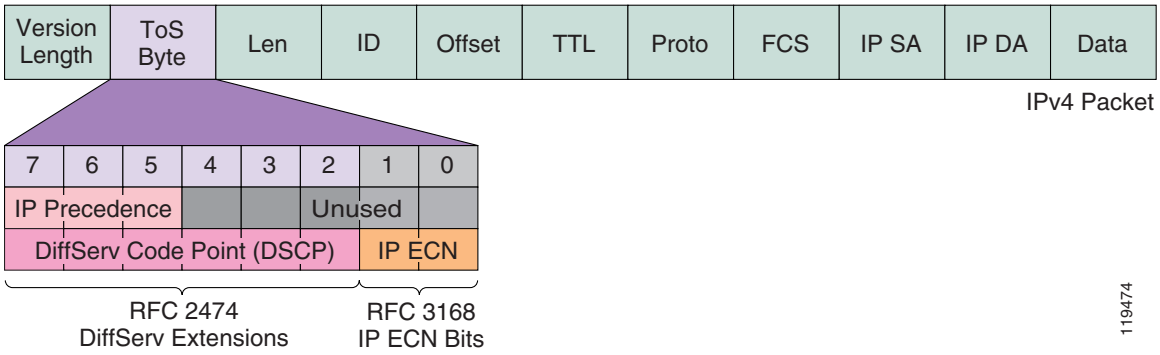
  - IPP values 6 and 7 are generally reserved for network control traffic such as routing.
  - IPP value 5 is recommended for voice.
  - IPP value 4 is shared by videoconferencing and streaming video.
  - IPP value 3 is for voice control.
  - IPP values 1 and 2 can be used for data applications.
  - IPP value 0 is the default marking value.

  Many enterprises find IPP marking to be overly restrictive and limiting, favoring instead the 6-Bit/64-value DSCP marking model.
- DSCPs and Per-Hop Behaviors (PHBs)—DSCP values can be expressed in numeric form or by special standards-based names called Per-Hop Behaviors. There are four broad classes of DSCP PHB markings: Best Effort (BE or DSCP 0), RFC 2474 Class Selectors (CS1–CS7, which are identical/backwards-compatible to IPP values 1–7), RFC 2597 Assured Forwarding PHBs (AF*xy*), and RFC 3268 Expedited Forwarding (EF).

  There are four Assured Forwarding classes, each of which begins with the letters “AF” followed by two numbers. The first number corresponds to the DiffServ Class of the AF group and can range from 1 through 4. The second number refers to the level of Drop Preference within each AF class and can range from 1 (lowest Drop Preference) through 3 (highest Drop Preference).

  DSCP values can be expressed in decimal form or with their PHB keywords. For example, DSCP EF is synonymous with DSCP 46, and DSCP AF31 is synonymous with DSCP 26.
- IP Explicit Congestion Notification (IP ECN)—IP ECN, as defined in RFC 3168, makes use of the last two bits of the IP ToS byte, which are not used by the 6-bit DSCP markings, as shown in Figure 1-2.

Figure 1-2 The IP ToS Byte (DSCP and IP ECN)



These last two bits are used to indicate to TCP senders whether or not congestion was experienced during transit. In this way, TCP senders can adjust their TCP windows so that they do not send more traffic than the network can service. Previously, dropping packets was the only way that congestion feedback could be signaled to TCP senders. Using IP ECN, however, congestion notification can be signaled without dropping packets. The first IP ECN bit (7th in the ToS byte) is used to indicate whether the device supports IP ECN and the second bit (last bit in the IP ToS byte) is used to indicate whether congestion was experienced (0="no congestion"; 1= "congestion was experienced"). IP ECN can be marked through a congestion avoidance mechanism such as weighted early random detection (WRED).

Policing and Markdown Tools

Policing tools (policers) determine whether packets are conforming to administratively-defined traffic rates and take action accordingly. Such action could include marking, remarking or dropping a packet.

A basic policer monitors a single rate: traffic equal to or below the defined rate is considered to *conform* to the rate, while traffic above the defined rate is considered to *exceed* the rate. On the other hand, the algorithm of a dual-rate policer (such as described in RFC 2698) is analogous to a traffic light. Traffic equal to or below the principal defined rate (green light) is considered to *conform* to the rate. An allowance for moderate amounts of traffic above this principal rate is permitted (yellow light) and such traffic is considered to *exceed* the rate. However, a clearly-defined upper-limit of tolerance is set (red light), beyond which traffic is considered to *violate* the rate.

Policers complement classification and marking policies. For example, as previously discussed, RFC 2597 defines the AF classes of PHBs. Traffic conforming to the defined rate of a given AF class is marked to the first Drop Preference level of a given AF class (for example, AF21). Traffic exceeding this rate is marked down to the second Drop Preference level (for example, AF22) and violating traffic is either marked down further to the third Drop Preference level (for example, AF23) or simply dropped.

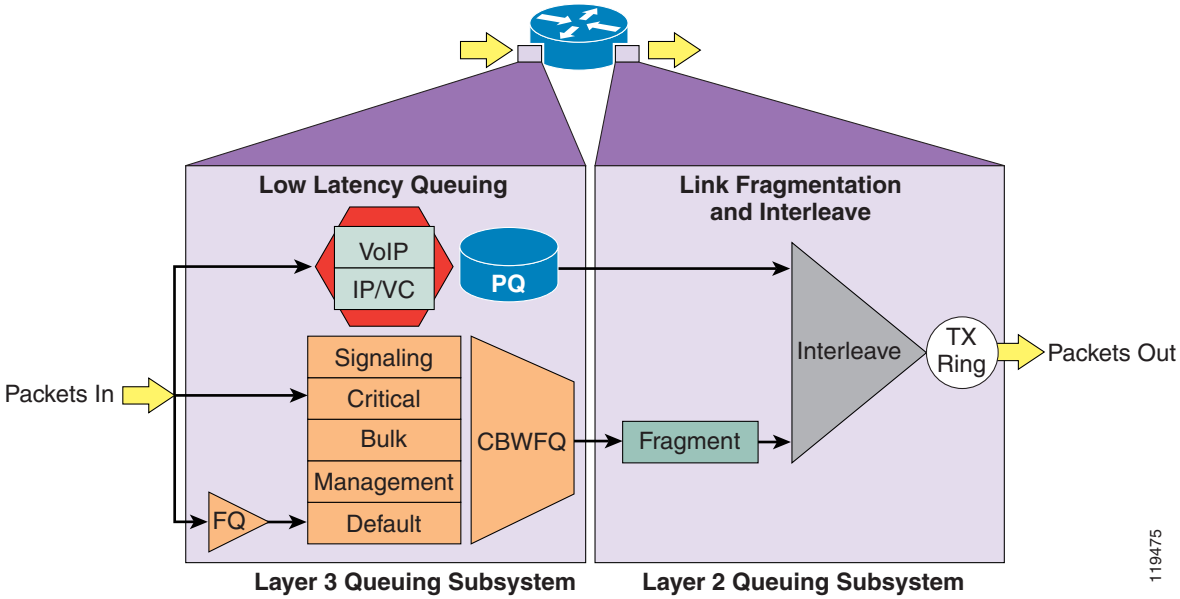
Scheduling Tools

Scheduling tools determine how a frame/packet exits a device. Whenever packets enter a device faster than they can exit it, such as with speed mismatches, then a point of congestion, or bottleneck, can occur. Devices have buffers that allow for scheduling higher-priority packets to exit sooner than lower priority ones, which is commonly called queueing.

Queueing algorithms are activated only when a device is experiencing congestion and are deactivated when the congestion clears. The main Cisco IOS software queuing tools are Low Latency Queueing (LLQ), which provides strict priority servicing and is intended for realtime applications such as VoIP; and Class-Based Weighted Fair Queuing (CBWFQ), which provides bandwidth guarantees to given classes of traffic and fairness to discrete traffic flows within these traffic classes.

Figure 1-3 shows the Layer 3 and Layer 2 queuing subsystems of the Cisco IOS software LLQ/CBWFQ algorithm.

Figure 1-3 LLQ/CBWFQ Operation



Queueing buffers act like a funnel for water being poured into a small opening. If water enters the funnel faster than it exits, eventually the funnel overflows from the top. When queueing buffers begin overflowing from the top, packets may be dropped either as they arrive (tail drop) or selectively before all buffers are filled.

Selective dropping of packets when the queues are filling is referred to as *congestion avoidance*. Congestion avoidance mechanisms work best with TCP-based applications because selective dropping of packets causes the TCP windowing mechanisms to "throttle-back" and adjust the rate of flows to manageable rates.

Congestion avoidance mechanisms are complementary to queueing algorithms. Queueing algorithms manage the front of a queue while congestion avoidance mechanisms manage the tail of the queue. Congestion avoidance mechanisms thus indirectly affect scheduling.

The principle IOS congestion avoidance mechanism is WRED, which randomly drops packets as queues fill to capacity. However, the randomness of this selection can be skewed by traffic weights. The weight can either be IP Precedence values, as is the case with default WRED which drops *lower* IPP values more aggressively (for example, IPP 1 would be dropped more aggressively than IPP 6) or the weights can be AF Drop Preference values, as is the case with DSCP-Based WRED which drops *higher* AF Drop Preference values more aggressively (for example, AF23 is dropped more aggressively than AF22, which in turn is dropped more aggressively than AF21). WRED can also be used to set the IP ECN bits to indicate that congestion was experienced in transit.

Link-Specific Tools

Link-specific tools include the following:

- Shaping tools—A shaper typically delays excess traffic above an administratively-defined rate using a buffer to hold packets and shape the flow when the data rate of the source is higher than expected.
- Link Fragmentation and Interleaving tools—With slow-speed WAN circuits, large data packets take an excessively long time to be placed onto the wire. This delay, called *serialization delay*, can easily cause a VoIP packet to exceed its delay and/or jitter threshold. There are two main tools to mitigate serialization delay on slow (768 kbps) links: Multilink PPP Link Fragmentation and Interleaving (MLP LFI) and Frame Relay Fragmentation (FRF.12).
- Compression tools—Compression techniques, such as compressed Real-Time Protocol (cRTP), minimize bandwidth requirements and are highly useful on slow links. At 40 bytes total, the header portion of a VoIP packet is relatively large and can account for nearly two-thirds or the entire VoIP packet (as in the case of G.729 VoIP). To avoid the unnecessary consumption of available bandwidth, you can use cRTP on a link-by-link basis. cRTP compresses IP/UDP/RTP headers from 40 bytes to between two and five bytes (which results in a bandwidth savings of approximately 66% for G.729 VoIP).
- Transmit ring (Tx-Ring) tuning—The Tx-Ring is a final interface First-In-First-Out (FIFO) queue that holds frames to be immediately transmitted by the physical interface. The Tx-Ring ensures that a frame is always available when the interface is ready to transmit traffic, so that link utilization is driven to 100 % of capacity. The size of the Tx-Ring is dependant on the hardware, software, Layer 2 media, and queueing algorithm configured on the interface. The Tx-Ring may have to be tuned on certain platforms/interfaces to prevent unnecessary delay/jitter introduced by this final FIFO queue.

AutoQoS Tools

The richness of the Cisco QoS toolset inevitably increases its deployment complexity. To address customer demand for simplification of QoS deployment, Cisco has developed the Automatic QoS (AutoQoS) features. AutoQoS is an intelligent macro that allows an administrator to enter one or two simple AutoQoS commands to enable all the appropriate features for the recommended QoS settings for an application on a specific interface.

AutoQoS VoIP, the first release of AutoQoS, provides best-practice QoS designs for VoIP on Cisco Catalyst switches and Cisco IOS routers. By entering one global and/or one interface command, depending on the platform, the AutoQoS VoIP macro expands these commands into the recommended VoIP QoS configurations (complete with all the calculated parameters and settings) for the platform and interface on which the AutoQoS is being applied.

For Campus Catalyst switches, AutoQoS automatically performs the following tasks:

- Enforces a trust boundary at Cisco IP Phones.
- Enforces a trust boundary on Catalyst switch access ports and uplinks/downlinks.
- Enables Catalyst strict priority queuing for voice and weighted round robin queuing for data traffic.
- Modifies queue admission criteria (CoS-to-queue mappings).
- Modifies queue sizes as well as queue weights where required.
- Modifies CoS-to-DSCP and IP Precedence-to-DSCP mappings.

For Cisco IOS routers, AutoQoS is supported on Frame Relay (FR), Asynchronous Transfer Mode (ATM), High-Level Data Link Control (HDLC), Point-to-Point Protocol (PPP), and FR-to-ATM links.

For Cisco IOS routers, AutoQoS automatically performs the following tasks:

- Classifies and marks VoIP bearer traffic (to DSCP EF) and Call-Signaling traffic (to DSCP CS3).
  - Applies scheduling:
  - Low Latency Queuing (LLQ) for voice
  - Class-Based Weighted Fair Queuing (CBWFQ) for Call-Signaling
  - Fair Queuing (FQ) for all other traffic
- Enables Frame Relay Traffic Shaping (FRTS) with optimal parameters, if required.
- Enables Link Fragmentation and Interleaving (LFI), either MLP LFI or FRF.12, on slow (768 kbps) links, if required.
- Enables IP RTP header compression (cRTP), if required.
- Provides Remote Monitoring (RMON) alerts of dropped VoIP packets.

AutoQoS VoIP became available on Cisco IOS router platforms in 12.2(15)T.

In its second release, for Cisco IOS routers only, AutoQoS Enterprise detects and provisions for up to ten classes of traffic, including the following:

- Voice
- Interactive-Video
- Streaming-Video
- Call-Signaling
- Transactional Data
- Bulk Data
- Routing
- Network Management
- Best Effort
- Scavenger

These classes will be explained in more detail later in this document.

The AutoQoS Enterprise feature consists of two configuration phases, completed in the following order:

- Auto Discovery (data collection)—Uses NBAR-based protocol discovery to detect the applications on the network and performs statistical analysis on the network traffic.
- AutoQoS template generation and installation—Generates templates from the data collected during the Auto Discovery phase and installs the templates on the interface. These templates are then used as the basis for creating the class maps and policy maps for your network. After the class maps and policy maps are created, they are then installed on the interface.

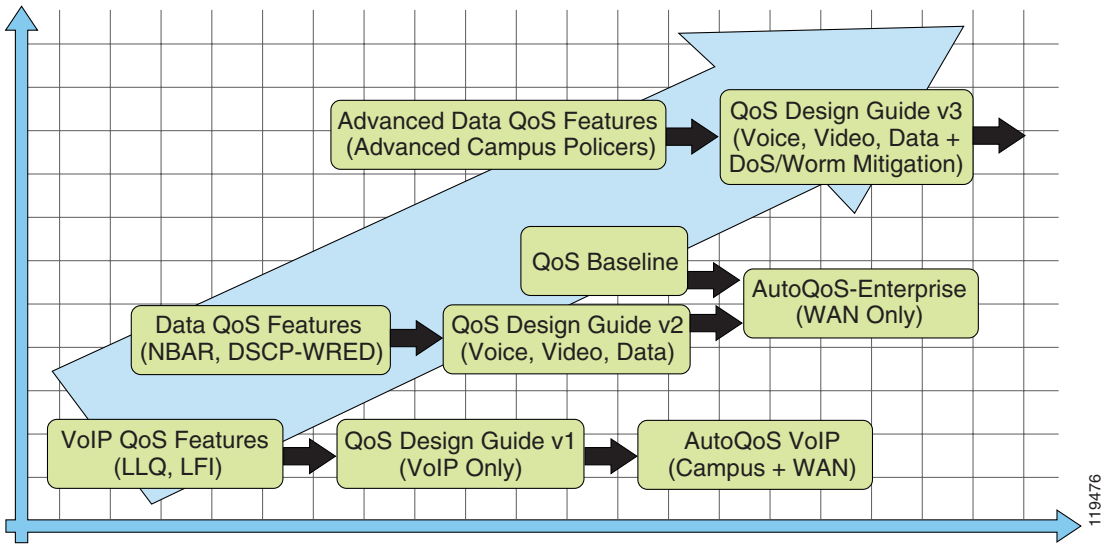
AutoQoS Enterprise became available on Cisco routers in Cisco IOS 12.3(7)T.

Some may naturally then ask: Why should I read the separate QoS design document when I have AutoQoS? While it is true that AutoQoS-VoIP is an excellent tool for customers with the objective of enabling QoS for VoIP (only) on their Campus and WAN infrastructures, and AutoQoS-Enterprise is a fine tool for enabling basic Branch-router WAN-Edge QoS for voice, video and multiple classes of data. For customers that have such basic QoS needs and don't have the time or desire to learn or do more with QoS, AutoQoS is definitely the way to go.

However, it is important to remember where AutoQoS came from. AutoQoS tools are the result of Cisco QoS feature development coupled with Cisco QoS Design Guides based on large-scale lab-testing. AutoQoS VoIP is the product of the first QoS Design Guide, one of the most popular/downloaded technical white papers ever produced within Cisco. AutoQoS Enterprise is the result of the strategic QoS Baseline (discussed later in this document) coupled with the second generation QoS Design Guide. These latest QoS design documents represents the third-generation QoS Design Guide, which is essentially a proposed blueprint for the next version of AutoQoS.

Figure 1-4 shows the relationship between Cisco QoS features, Design Guides, and AutoQoS.

Figure 1-4 Cisco QoS Feature, Design Guide and AutoQoS Evolution



Call Admission Control Tools

After performing the calculations to provision the network with the required bandwidth to support voice, video and data applications, you must ensure that voice or video do not oversubscribe the portion of the bandwidth allocated to them. While most DiffServ QoS features are used to protect voice from data, Call Admission Control (CAC) tools are used to protect voice from voice and video from video.

CAC tools fall into the following three main categories:

- Local—Local CAC mechanisms are a voice gateway router function, typically deployed on the outgoing gateway. The CAC decision is based on nodal information such as the state of the outgoing LAN/WAN link that the voice call traverses if allowed to proceed. Local mechanisms include configuration items to disallow more than a fixed number of calls.

  If the network designer already knows that no more than five VoIP calls will fit across the outgoing WAN link's LLQ configuration because of bandwidth limitations, then it would be recommended to configure the local gateway node to not allow more than five simultaneous calls.

- Measurement-based—Measurement-based CAC techniques look ahead into the packet network to gauge the state of the network to determine whether or not to allow a new call. This usually implies sending probes to the destination IP address, which could be the terminating gateway or endpoint, or another device in between.

The probes return to the outgoing gateway or endpoint information on the conditions found while traversing the network to the destination. Typically, loss and delay characteristics are the interesting elements of information for voice CAC decisions. The outgoing device then uses this information in combination with configured information to decide if the network conditions exceed a given or configured threshold.

- Resource-based—There are two types of resource-based mechanisms: those that calculate resources needed and/or available, and those that reserve resources for the call. Resources of interest include link bandwidth, DSPs and DS0 timeslots on the connecting TDM trunks to a voice gateway, CPU power and memory. Several of these resources could be constrained at one or more nodes that the call traverses to its destination.

Cisco CallManager has additional CAC features to handle management of VoIP network deployments. These features are not mutually exclusive to the features listed above. While CallManager Location-Based CAC is deployed in the overall network to manage VoIP bandwidth availability for both Cisco IP Phones and voice gateways, local measurement-based or resource-based features may be deployed at the same time on the voice gateway to push back calls into the private Branch exchange (PBX) or publicly-switched telephone network (PSTN) if IP network conditions do not allow their entry into the VoIP network.

Note A detailed discussion of CAC configuration is beyond the scope of this document, but CAC configuration is crucial for a successful VoIP deployment. For additional information on CallManager CAC, refer to the IP Telephony Solution Reference Network Design Guide at http://www.cisco.com/en/US/products/sw/voicesw/ps556/products_implementation_design_guides_list.html.

How is QoS Optimally Deployed within the Enterprise?

A successful QoS deployment is comprised of multiple phases, including:

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.

Each of these phases will be addressed in more detail in the following sections.

1) Strategically Defining QoS Objectives

QoS technologies are the enablers for business/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, among the first questions that arise during a QoS deployment are: How many traffic classes should be provisioned for? And what should they be?

To help answer these fundamental questions, organizational objectives need to be defined, such as:

- Is the objective to enable VoIP only or video also required?
- If so, is video-conferencing required or streaming video? Or both?

- Are there applications that are considered mission-critical, and if so, what are they?
- Does the organization wish to squelch certain types of traffic, and if so, what are they?

To help address these crucial questions and to simplify QoS, Cisco has adopted a new initiative called the "QoS Baseline." The QoS Baseline is a strategic document designed to unify QoS within Cisco, from enterprise to service provider and from engineering to marketing. The QoS Baseline was written by Cisco's most qualified QoS experts, who have developed or contributed to the related IETF RFC standards (as well as other technology standards) and are thus eminently qualified to interpret these standards. The QoS Baseline also provides uniform, standards-based recommendations to help ensure that QoS designs and deployments are unified and consistent.

The QoS Baseline defines up to 11 classes of traffic that may be viewed as critical to a given enterprise. A summary these classes and their respective standards-based marking recommendations are presented in Table 1-1.

Table 1-1 Cisco QoS Baseline/Technical Marketing (Interim) Classification and Marking Recommendations

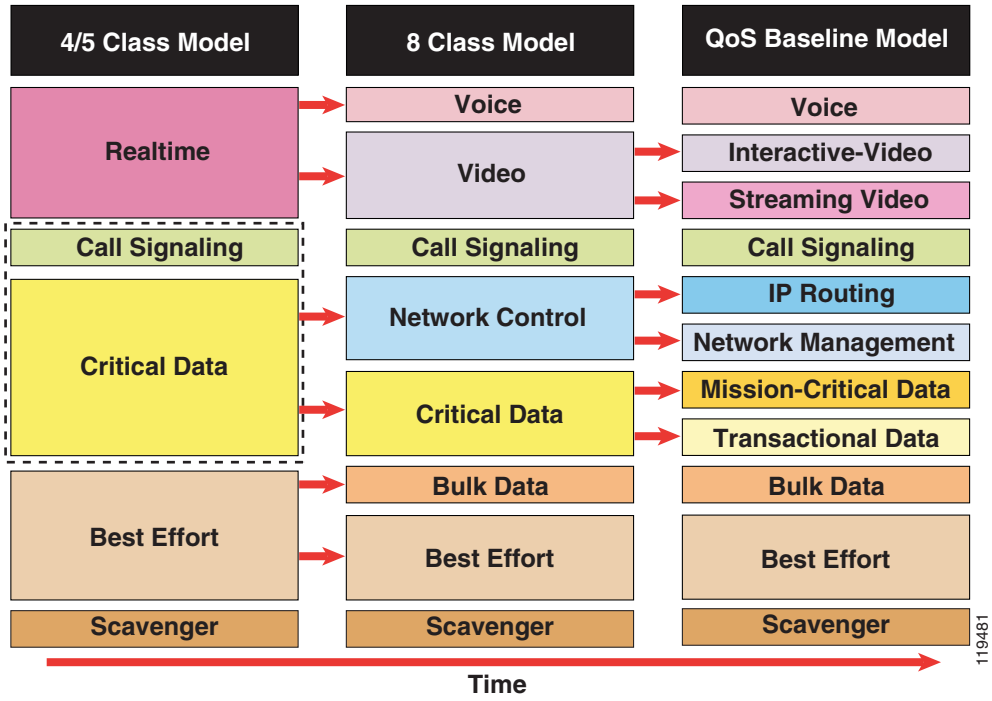
Application	Layer 3 Classification			Layer 2 CoS/MPLS EXP
	IPP	PHB	DSCP	
IP Routing	6	CS6	48	6
Voice	5	EF	46	5
Interactive Video	4	AF41	34	4
Streaming-Video	4	CS4	32	4
Locally-Defined Mission-Critical Data (see note below)	3	—	25	3
Call-Signaling (see note below)	3	AF31/CS3	26/24	3
Transactional Data	2	AF21	18	2
Network Management	2	CS2	16	2
Bulk Data	1	AF11	10	1
Scavenger	1	CS1	8	1
Best Effort	0	0	0	0

Note The QoS Baseline recommends marking Call-Signaling to CS3. However, currently most Cisco IP Telephony products mark Call-Signaling to AF31. A marking migration from AF31 to CS3 is under way within Cisco, but in the interim it is recommended that both AF31 and CS3 be reserved for Call-Signaling and that Locally-Defined Mission-Critical Data applications be marked to a temporary placeholder non-standard DSCP, such as 25. Upon completion of the migration, the QoS Baseline marking recommendations of CS3 for Call-Signaling and AF31 for Locally-Defined Mission-Critical Data applications should be used. These marking recommendations are more in line with RFC 2474 and RFC 2597.

Enterprises do not need to deploy all 11 classes of the QoS Baseline 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 their 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 Figure 1-5.

Figure 1-5 Example Strategy for Expanding the Number of Classes of Service over Time



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, garnishing executive endorsement whenever possible.

A strategic standards-based guide like the QoS Baseline 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.

2) Analyzing Application Service-Level Requirements

The following sections present an overview of the QoS requirements for voice, video and multiple classes of data, including the following topics:

- QoS Requirements of VoIP
- QoS Requirements of Video
- QoS Requirements of Data Applications
- QoS Requirements of the Control Plane
- QoS Requirements of the Scavenger Class

QoS Requirements of VoIP

This section includes the following topics:

- Voice (Bearer Traffic)
- Call-Signaling Traffic

VoIP deployments require provisioning explicit priority servicing for VoIP (bearer stream) traffic and a guaranteed bandwidth service for Call-Signaling traffic. These related classes will be examined separately.

Voice (Bearer Traffic)

A summary of the key QoS requirements and recommendations for Voice (bearer traffic) are:

- **Voice** traffic should be marked to **DSCP EF** per the QoS Baseline and RFC 3246.
- **Loss** should be no more than **1 %.**
- One-way **Latency** (mouth-to-ear) should be no more than **150 ms**.
- Average one-way **Jitter** should be targeted under **30 ms**.
- **21–320 kbps of guaranteed priority bandwidth is required per call** (depending on the sampling rate, VoIP codec and Layer 2 media overhead).

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

Loss causes voice clipping and skips. The packetization interval determines the size of samples contained within a single packet. Assuming a 20 ms (default) packetization interval, the loss of two or more consecutive packets results in a noticeable degradation of voice quality. 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.

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).

If the end-to-end voice delay becomes too long, the conversation begins to sound like two parties talking over a satellite link or even a CB radio. While the ITU G.114 states that a 150 ms one-way (mouth-to-ear) 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, then the delay boundary can be extended to 200 ms without significant impact on voice quality.

Note Higher delays may also be viewed as acceptable to certain organizations, but the corresponding reduction in VoIP quality must be taken into account when making such design decisions.

Jitter buffers (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 set either arbitrarily large or arbitrarily small, then it imposes unnecessary constraints on the characteristics of the network. A jitter buffer set too large adds to the end-to-end delay, meaning that less delay budget is available for the network such that the network needs to support a delay target tighter than practically necessary. If a jitter buffer is too small to accommodate the network jitter, then buffer underflows or overflows can occur.

An underflow is when the buffer is empty when the codec needs to play out a sample. An overflow is when the jitter buffer is already full and another packet arrives that cannot therefore 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. Well-designed adaptive jitter buffer algorithms should not impose any unnecessary constraints on the network design by:

Instantly increasing the jitter buffer size to the current measured jitter value following a jitter buffer overflow.

Slowly decreasing the jitter buffer size when the measured jitter is less that the current jitter buffer size.

Using a Programmable Logic Controller (PLC) to interpolate for the loss of a packet on a jitter buffer underflow.

Where such adaptive jitter buffers are used, we can in theory engineer out explicit considerations of jitter by accounting for worst-case per hop delays. Advanced formulas can be used to arrive at network-specific design recommendations for jitter based on maximum and minimum per-hop delays. Alternatively, this 30 ms value can be used as a jitter target as 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 bandwidth consumed by VoIP streams (in bps) 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 8 (to convert it to bits) and then multiplying again by the packetization rate (default of 50 packets per second).

Table 1-2 details the bandwidth per VoIP flow at a default packet rate of 50 packets per second (pps). This does not include Layer 2 overhead and does not take into account any possible compression schemes, such as cRTP.

Table 1-2 Voice Bandwidth (without Layer 2 Overhead)

Bandwidth Consumption	Sampling Rate	Voice Payload in Bytes	Packets per Second	Bandwidth per Conversation
G.711	20 ms	160	50	80 kbps

Table 1-2 Voice Bandwidth (without Layer 2 Overhead)

G.711	30 ms	240	33	74 kbps
G.729A	20 ms	20	50	24 kbps
G.729A	30 ms	30	33	19 kbps

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.

Table 1-3 shows a more accurate bandwidth provisioning example for voice because it includes Layer 2 overhead.

Table 1-3 Voice Bandwidth (Including Layer 2 Overhead)

Bandwidth Consumption	802.1Q Ethernet	PPP	MLP	Frame-Relay w/FRF.12	ATM
G.711 at 50 pps	93 kbps	84 kbps	86 kbps	84 kbps	106 kbps
G.711 at 33 pps	83 kbps	77 kbps	78 kbps	77 kbps	84 kbps
G.729A at 50 pps	37 kbps	28 kbps	30 kbps	28 kbps	43 kbps
G.729A at 33 pps	27 kbps	21 kbps	22 kbps	21 kbps	28 kbps

Note A handy tool for quickly and accurately calculating VoIP bandwidth requirements (factoring in the codec, the use of cRTP and L2 overhead) can be found at: http://tools.cisco.com/Support/VBC/jsp/Codec_Calc1.jsp

Call-Signaling Traffic

The following are key QoS requirements and recommendations for Call-Signaling traffic:

- **Call-Signaling** traffic should be marked as **DSCP CS3** per the QoS Baseline (during migration, it may also be marked the legacy value of DSCP AF31).
- **150 bps** (plus Layer 2 overhead) per phone of **guaranteed bandwidth** is required for voice control traffic; more may be required, depending on the call signaling protocol(s) 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 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/aggressive dropping. Some Cisco IP Telephony products have already begun transitioning to DSCP CS3 for Call-Signaling marking. In this interim period, both code-points (CS3 and AF31) should be reserved for Call-Signaling marking until the transition is complete.

- Many Cisco IP phones use Skinny Call-Control Protocol (SCCP) for call signaling. SCCP is a relatively lightweight protocol that requires only a minimal amount of bandwidth protection. However, newer versions of CallManager and SCCP have improved functionality requiring new message sets yielding a higher bandwidth consumption. Cisco signaling bandwidth design recommendations have been adjusted to match. The IPT SRND's Network Infrastructure chapter contains the relevant details, available at: http://www.cisco.com/en/US/products/sw/voicesw/ps556/products_implementation_design_guides_list.html.
- Other call signaling protocols include (but are not limited to) H.323, H.225, Session Initiated Protocol (SIP) and Media Gateway Control Protocol (MGCP). Each call signaling protocol has unique TCP/UDP ports and traffic patterns that should be taken into account when provisioning QoS policies for them.

QoS Requirements of Video

This section describes the two main types of video traffic, and includes the following topics:

- Interactive Video
- Streaming Video

Interactive Video

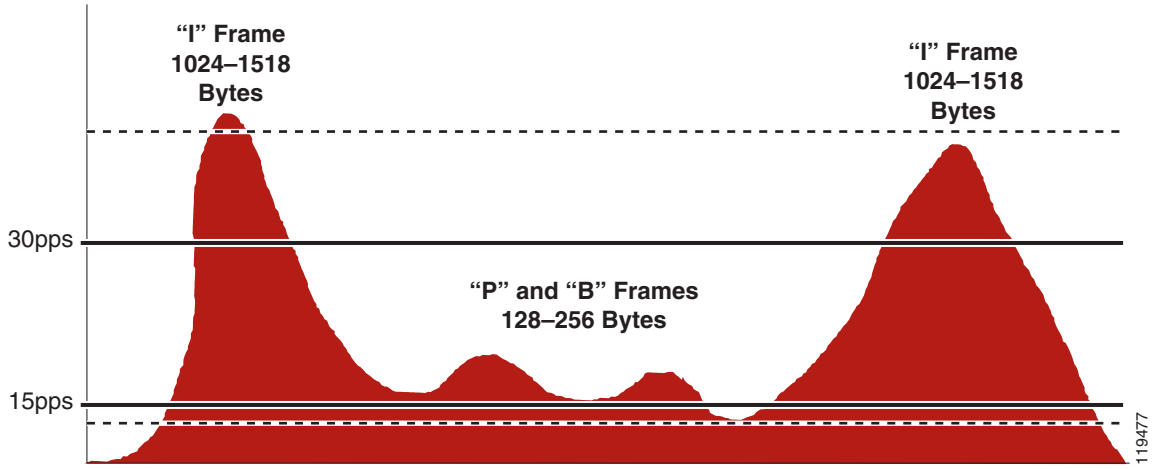
When provisioning for Interactive Video (IP Videoconferencing) traffic, the following guidelines are recommended:

- **Interactive Video** traffic should be marked to **DSCP AF41**; excess Interactive-Video traffic can be marked down by a policer to AF42 or AF43.
- **Loss** should be no more than **1 %.**
- One-way **Latency** should be no more than **150 ms**.
- **Jitter** should be no more than **30 ms**.
- **Overprovision Interactive Video queues by 20%** to accommodate bursts

Because IP Videoconferencing (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 videoconferencing are radically different from voice.

For example, videoconferencing traffic has varying packet sizes and extremely variable packet rates, as shown in Figure 1-6.

Figure 1-6 IP Videoconferencing Traffic Rates and Packet Sizes



The videoconferencing rate is the sampling rate of the video stream, not the actual bandwidth the video call requires. In other words, the data payload of videoconferencing packets is filled with 384 kbps 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 a conservative rule of thumb for IP/VC bandwidth provisioning is to overprovision the guaranteed/priority bandwidth by 20 percent. For example, a 384 kbps IP/VC stream would be adequately provisioned with an LLQ/CBWFQ of 460 kbps.

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 Videoconferencing (IP/VC) stream. For multiple streams, this burst parameter may be increased as required.

Streaming Video

When addressing the QoS needs of Streaming Video traffic, the following guidelines are recommended:

- **Streaming Video** (whether unicast or multicast) should be marked to **DSCP CS4** as designated by the QoS Baseline.
- **Loss** should be no more than **5 %**.
- **Latency** should be no more than **4–5 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.
- Streaming video is typically **unidirectional** and, therefore, Branch routers may not require provisioning for Streaming Video traffic on their WAN/VPN edges (in the direction of Branch-to-Campus).

- Non-organizational Streaming Video applications, such as entertainment videos, may be marked as Scavenger (DSCP CS1) and assigned a minimal bandwidth (CBWFQ) percentage. For more information, see Scavenger-class QoS DoS/Worm Mitigation Strategy.

Streaming Video applications have more lenient QoS requirements because they are delay-insensitive (the video can take several seconds to cue-up) and are largely 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.

The QoS Baseline recommendation for Streaming Video marking is DSCP CS4.

An interesting consideration with respect to Streaming Video comes into play when designing WAN/VPN edge policies on Branch routers: because Streaming Video is generally unidirectional, a separate class would likely not be needed for this traffic class in the Branch-to-Campus direction of traffic flow.

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

QoS Requirements of Data Applications

This section includes the following topics:

- Best Effort Data
- Bulk Data
- Transactional/Interactive Data
- Locally-Defined Mission-Critical Data

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, and so on. Not only do applications vary one from another, but even the same application can vary significantly in one *version* to another.

Given this, how best to provision QoS for data is a daunting question. The Cisco QoS Baseline identifies four main classes of data traffic, according to their general networking characteristics and requirements. These classes are Best Effort, Bulk Data, Transactional/Interactive Data and Locally-Defined Mission-Critical Data.

Best Effort Data

The Best Effort class is the default class for all data traffic. An application will be removed from the default class only if it has been selected for preferential or deferential treatment.

When addressing the QoS needs of Best Effort data traffic, Cisco recommends the following guidelines:

- **Best Effort** traffic should be marked to **DSCP 0**.
- Adequate bandwidth should be assigned to the Best Effort class as a whole, because the majority of applications will default to this class; reserve **at least 25 percent for Best Effort traffic**.

In 2003, a Wall Street financial company did an extensive study to identify and categorize the number of different applications on their networks. They found over 3000 discrete applications traversing their infrastructure. Further research has shown that this is not uncommon for larger enterprises.

Because enterprises have several hundred, if not thousands, of data applications running over their networks (of which, the majority will default to the Best Effort class), you need to provision adequate bandwidth for the default class as a whole, to handle the sheer volume of applications that will be included in it. Otherwise, applications defaulting to this class will be easily drowned out, which typically results in an increased number of calls to the networking help desk from frustrated users. Cisco therefore recommends that you reserve at least 25 percent of link bandwidth for the default Best Effort class.

Bulk Data

The Bulk 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 the following:

- FTP
- E-mail
- Backup operations
- Database synchronizing or replicating operations
- Content distribution
- Any other type of background operation

When addressing the QoS needs of 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).
- Bulk Data traffic should have a **moderate bandwidth guarantee**, but should be **constrained** from dominating a link.

The advantage of provisioning moderate bandwidth guarantees to Bulk Data applications rather than applying policers to them is that Bulk applications can dynamically take advantage of unused bandwidth and thus speed up their operations during non-peak periods. This in turn reduces the likelihood of their bleeding into busy periods and absorbing inordinate amounts of bandwidth for their time-insensitive operations.

Transactional/Interactive Data

The Transactional/Interactive Data class, also referred to simply as Transactional Data, is a combination to 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.

E-mail is not considered a Transactional Data client-server application, as most e-mail operations occur in the background and users do not usually notice even several hundred millisecond delays in mailspool operations.

When addressing the QoS needs of 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 they support.

Locally-Defined Mission-Critical Data

The Locally-Defined Mission-Critical Data class is probably the most misunderstood class specified in the QoS Baseline. Under the QoS Baseline model, all traffic classes (with the exclusion of Scavenger and Best Effort) are considered critical to the enterprise. The term "locally-defined" is used to underscore the purpose of this class, which is to provide each enterprise with a premium class of service for a select subset of their Transactional Data applications that have the highest business priority for them.

For example, an enterprise may have properly provisioned Oracle, SAP, BEA, and DLSw+ within their Transactional Data class. However, the majority of their revenue may come from SAP, and therefore they may want to give this Transactional Data application an even higher level of preference by assigning it to a dedicated class such as the Locally-Defined Mission-Critical Data class.

Because the admission criteria for this class is non-technical (being determined by business relevance and organizational objectives), the decision of which applications should be assigned to this special class can easily become an organizationally- and politically-charged debate. Cisco recommends that you assign as few applications to this class from the Transactional Data class as possible. You should also obtain executive endorsement for application assignments to the Locally-Defined Mission-Critical Data class, because the potential for QoS deployment derailment exists without such an endorsement.

For the sake of simplicity, this class will be referred to simply as Mission-Critical Data.

When addressing the QoS needs of Mission-Critical Data traffic, Cisco recommends the following guidelines:

- **Mission-Critical Data** traffic should be marked to **DSCP AF31**; excess mission-critical data traffic can then be marked down by a policer to AF32 or AF33. However, DSCP AF31 is currently being used by Cisco IP Telephony equipment as Call-Signaling, so until all Cisco IPT products mark Call-Signaling to DSCP CS3, a **temporary placeholder code point (DSCP 25)** can be used to identify Mission-Critical Data traffic.
- Mission-Critical Data traffic should have an **adequate bandwidth guarantee** for the interactive, foreground operations they support.

Table 1-4 shows some applications and the generic networking characteristics that determine for which data application class they are best suited.

Table 1-4 Data Applications by Class

Application Class	Example Applications	Application/Traffic Properties	Packet / Message Sizes
Interactive	Telnet, Citrix, Oracle Thin-Clients AOL Instant Messenger Yahoo Instant Messenger PlaceWare (Conference) Netmeeting Whiteboard	Highly-interactive applications with tight user feedback requirements.	Average message size < 100 B Max message size < 1 KB

Table 1-4 Data Applications by Class

Transactional	SAP, PeopleSoft (Vantive) Oracle—financials, Internet procurement, B2B, supply chain management, and application server Oracle 8i Database Ariba Buyer I2, Siebel, E.piphany Broadvision IBM Bus 2 Bus Microsoft SQL BEA Systems DLSw+	Transactional applications 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. Query response may consist of many TCP and FTP sessions running simultaneously (for example, HTTP based applications)	Depends on application—could be anywhere from 1 KB to 50 MB
Bulk	Database syncs Network-based backups Lotus Notes, Microsoft Outlook E-mail download (SMTP, POP3, IMAP, Exchange) Video content distribution, Large ftp file transfers	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		

QoS Requirements of the Control Plane

This section includes the following topics:

- IP Routing
- Network Management

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.

IP Routing

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.

Note that *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 traffic, Cisco recommends the following guidelines:

- **IP Routing** traffic should be marked to **DSCP CS6**; this is default behavior on Cisco IOS platforms.
- IGPs are usually adequately protected with the Cisco IOS internal PAK_Priority mechanism; Cisco recommends that EGPs such as BGP have an **explicit class for IP routing with a minimal bandwidth guarantee**.
- Cisco IOS automatically marks IP Routing traffic to DSCP CS6.

Additional information on PAK_PRIORITY can be found at:
http://www.cisco.com/warp/public/105/rtgupdates.html

Network Management

When addressing the QoS needs of Network Management traffic, Cisco recommends the following guidelines:

- **Network Management** traffic should be marked to **DSCP CS2**.
- Network Management applications should be explicitly protected with a **minimal bandwidth guarantee**.

Network management traffic is important to perform trend and capacity analysis and troubleshooting. Therefore, you can provision a separate minimal bandwidth queue for Network Management traffic, which could include SNMP, NTP, Syslog, NFS and other management applications.

QoS Requirements of the Scavenger Class

The Scavenger class, based on an Internet-II draft, is intended to provide deferential services, or “less-than-Best-Effort” services, to certain applications.

Applications assigned to this class have little or no contribution to the organizational objectives of the enterprise and are typically entertainment-oriented. These include: Peer-to-Peer (P2P) media-sharing applications (such as KaZaa, Morpheus, Grokster, Napster, iMesh, and so on), gaming applications (Doom, Quake, Unreal Tournament, and so on), and any entertainment video applications.

Assigning a minimal bandwidth queue to Scavenger traffic forces it to be squelched to virtually nothing during periods of congestion, but allows it to be available if bandwidth is not being used for business purposes, such as might occur during off-peak hours. This allows for a flexible, non-stringent policy control of non-business applications.

When provisioning for Scavenger traffic, Cisco recommends the following guidelines:

- **Scavenger** traffic should be marked to **DSCP CS1**.
- Scavenger traffic should be assigned the **lowest configurable queuing service**; for instance, in Cisco IOS this would mean assigning a **CBWFQ of 1 %** to Scavenger.

The Scavenger class is a critical component to the DoS/worm mitigation strategy presented later in this document.

3) Designing the QoS Policies

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. A separate QoS design document delves into the specific details of LAN, WAN, and VPN (both MPLS and IPSec VPN) QoS designs. 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 as such 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:

- Classification and Marking Principles
- Policing and Markdown Principles
- Queueing and Dropping Principles

Classification and Marking Principles

When classifying and marking traffic, an unofficial Differentiated Services 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. 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 received 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-realtime traffic. Such abuse could easily ruin the service quality of realtime 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 0–7), as does MPLS EXP. Therefore, only up to 8 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 QoS Baseline marking recommendations are standards-based, enterprises can easily adopt these markings to interface with service provider classes of service. Network mergers—whether the result of acquisitions, mergers or strategic-alliances—are also easier to manage when you use standards-based DSCP markings.

Policing and Markdown Principles

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/worm-generated traffic can masquerade under legitimate, well-known TCP/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 AF*x*1 should be marked down to AF*x*2 (or AF*x*3, whenever dual-rate policing—such as 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, and so this standards-based strategy cannot be implemented fully at this time. As an alternative workaround, single-rate policers can be configured to markdown excess traffic to DSCP CS1 (Scavenger); dual-rate policers can be configured to mark down excess traffic to AF*x*2, 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, then you should mark down Assured Forwarding classes by RFC 2597 rules to comply more closely with this standard.

Queuing and Dropping Principles

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/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.

When provisioning queuing, some best practice rules of thumb 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/deferential treatment is it removed from the default class. Because many enterprises have several hundred, if not thousands, of data applications 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 "Realtime" or "Strict Priority" class (which corresponds to RFC 3246 "An Expedited Forwarding Per-Hop Behavior"). The amount of bandwidth assigned to the Realtime queuing class is variable. However, if you assign too much traffic for strict priority queuing, then the overall effect is a dampening of QoS functionality for non-realtime applications. Remember: the goal of convergence is to enable voice, video, and data to *transparently* co-exist on a single network. When Realtime applications such as Voice or Interactive-Video dominate a link (especially a WAN/VPN link), then 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 realtime 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 realtime applications with data applications.

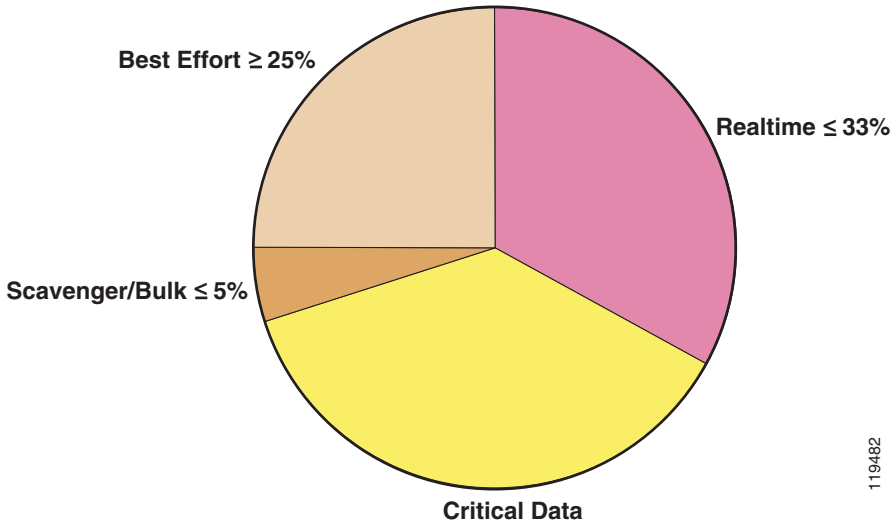
Cisco IOS software allows the abstraction (and thus configuration) of multiple strict priority LLQs. In such a multiple LLQ context, this design principle would *apply to the sum of all LLQs to be within one-third of link capacity.*

Note 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-realtime-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/Bulk queuing class a bandwidth percentage of 5. If you can uniquely assign Scavenger and Bulk Data to different queues, then you should assign the Scavenger queue a bandwidth percentage of 1.

The Realtime, Best Effort and Scavenger queuing best practice principles are shown in Figure 1-7.

Figure 1-7 Realtime, Best Effort and Scavenger Queuing Rules



Because platforms support a variety of queuing structures, configure consistent queuing policies according to platform capabilities to ensure consistent PHBs.

For example, on a platform that only supports four queues with CoS-based admission (such as a Catalyst switch) a basic queuing policy could be as follows:

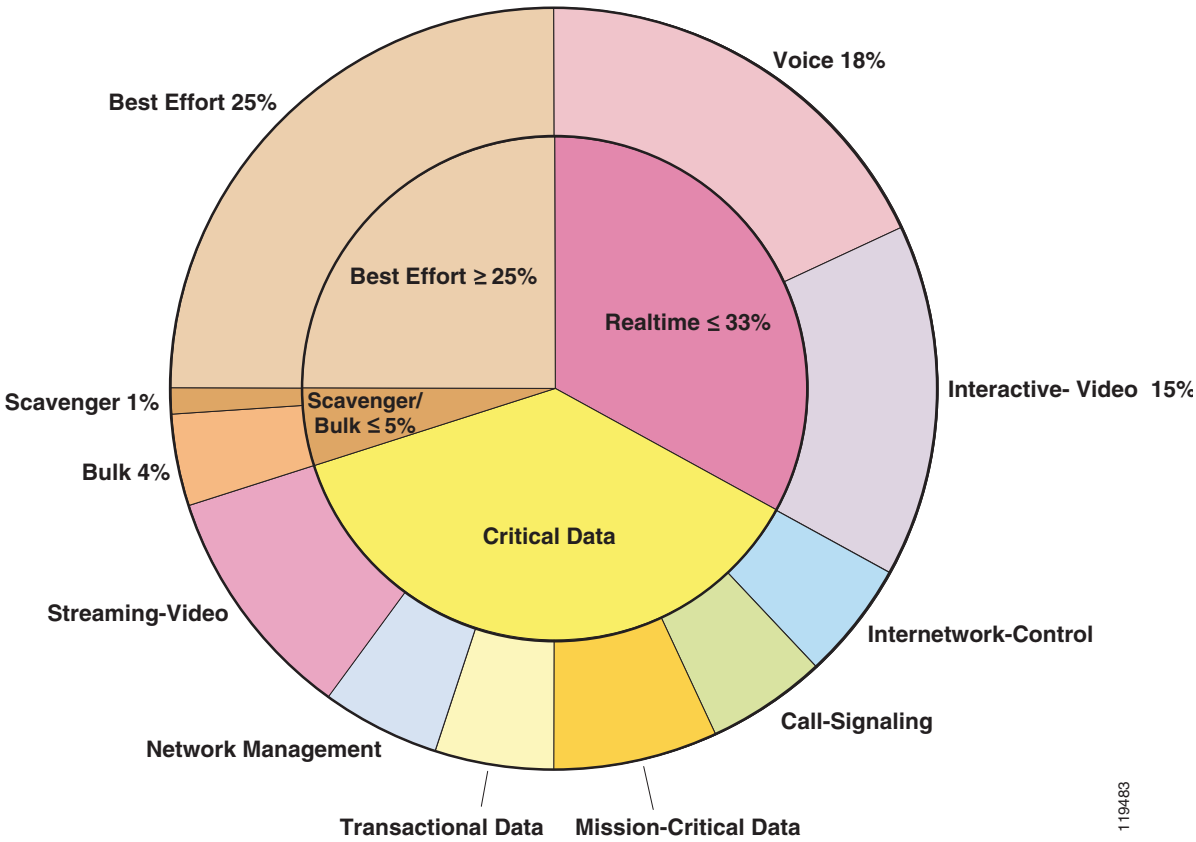
- Realtime (33%)
- Critical Data
- Best Effort Data (25%)
- Scavenger/Bulk (5%)

The queuing policies can be expanded on a platform that supports a full 11-class QoS Baseline queuing model in such a way as to provide consistent servicing to Realtime, Best Effort and Scavenger traffic. For example, on a platform such as a Cisco IOS router that supports 11 queues with DSCP-based admission, an advanced queuing policy could be as follows:

- Voice (18%)
- Interactive Video (15%)
- Internetwork-Control
- Call-Signaling
- Mission-Critical Data
- Transactional Data
- Network Management
- Streaming Video
- Best Effort Data (25%)
- Bulk Data (4%)
- Scavenger (1%)

The inter-relationship between these compatible queuing models is shown in Figure 1-8.

Figure 1-8 Compatible Four-Class and Eleven-Class Queuing Models following Realtime, Best Effort and Scavenger Queuing Rules



In this way, traffic receives compatible queuing at each node, regardless of platform capabilities, which is the overall objective of DiffServ PHB definitions.

Whenever supported, you should ***enable WRED (preferably DSCP-based WRED)*** on all TCP flows. WRED congestion avoidance prevents TCP global synchronization and increases overall throughput and link efficiency. Enabling WRED on UDP flows is optional.

These and other architecture-specific QoS design best-practices are discussed in more detail in a separate QoS design document, along with the configuration examples.

Furthermore, it is highly-recommended to schedule Proof-of-Concept (PoC) tests to verify that the hardware/software platforms in production support the required QoS features *in combination* with all the other features they are currently running. Remember, ***in theory, theory and practice are the same***. In other words, ***there is no substitute for testing***.

4) Rolling out the QoS Policies

Once the QoS designs have been finalized and PoC tested, it is vital to ensure that the networking team ***thoroughly understand the QoS features and syntax before enabling features on production networks.*** Such knowledge is critical for both rollout and subsequent troubleshooting of QoS-related issues.

Furthermore, it is recommended to ***schedule network downtime in order to rollout QoS*** features. While QoS is required end-to-end, it does not have to be deployed end-to-end at a single instance. A pilot network-segment can be selected for an initial deployment, and pending observation, the ***rollout can be expanded in stages*** to encompass the entire enterprise.

A rollback strategy is always recommended, to address unexpected issues arising from the QoS deployment.

5) Monitoring the Service-Levels

Implementing a QoS solution is not a one-time task that is complete upon policy deployment. A successful QoS policy ***rollout is followed by ongoing monitoring of service levels and periodic adjustments and tuning*** of QoS policies.

Short-term monitoring is useful for verifying that the deployed QoS policies are having the desired end-to-end effect. *Long-term monitoring* (trending) is needed to determine whether the provisioned bandwidth is still adequate for the changing needs of the enterprise. For example, upgrading to a newer version of an application may cause the provisioned bandwidth to be exceeded, as would the addition of new users. Furthermore, business objectives or economic climates themselves may change, and periodically the overall ranking of priority of applications may need revision.

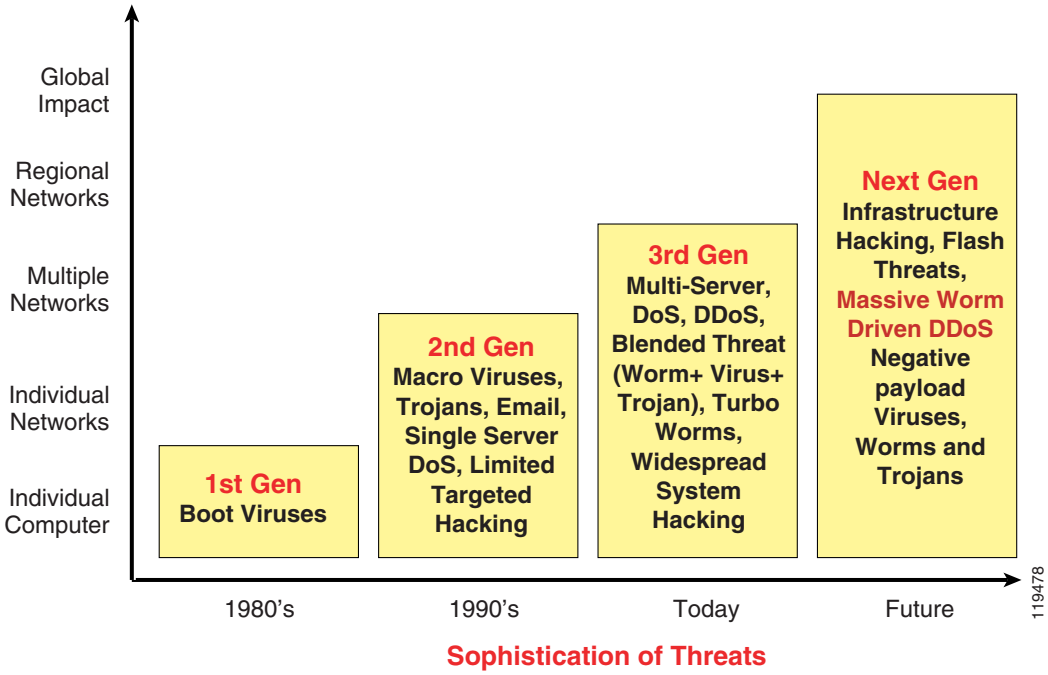
As business conditions change, the enterprise may need to adapt to these changes and ***may be required to begin the QoS deployment cycle anew***, by redefining their objectives, tuning and testing corresponding designs, rolling these new designs out and monitoring them to see if they match the redefined objectives.

How Can I Use QoS Tools to Mitigate DoS/Worm Attacks?

Whenever the business objectives of the enterprise includes mitigating DoS/worm attacks, the Scavenger-class QoS strategy and best practices described in this section apply.

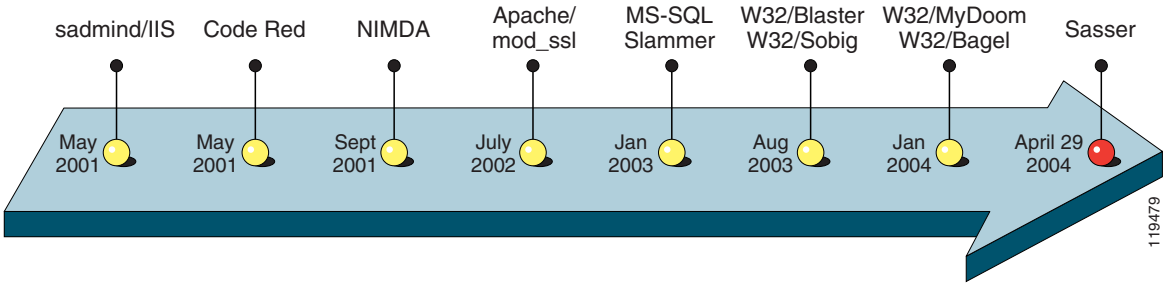
Worms have existed in one form or another since the beginning of the Internet, and have steadily increased in complexity and scope of damage, as shown in Figure 1-9.

Figure 1-9 Business Security Threat Evolution



There has been an exponential increase since 2001 in not only the frequency of DoS/worm attacks, but also in their relative sophistication. For example, more than 994 new Win32 viruses and worms were documented in the first half of 2003, more than double the 445 documented in the first half of 2002. Some of these more recent worms are shown in Figure 1-10.

Figure 1-10 Recent Internet Worms



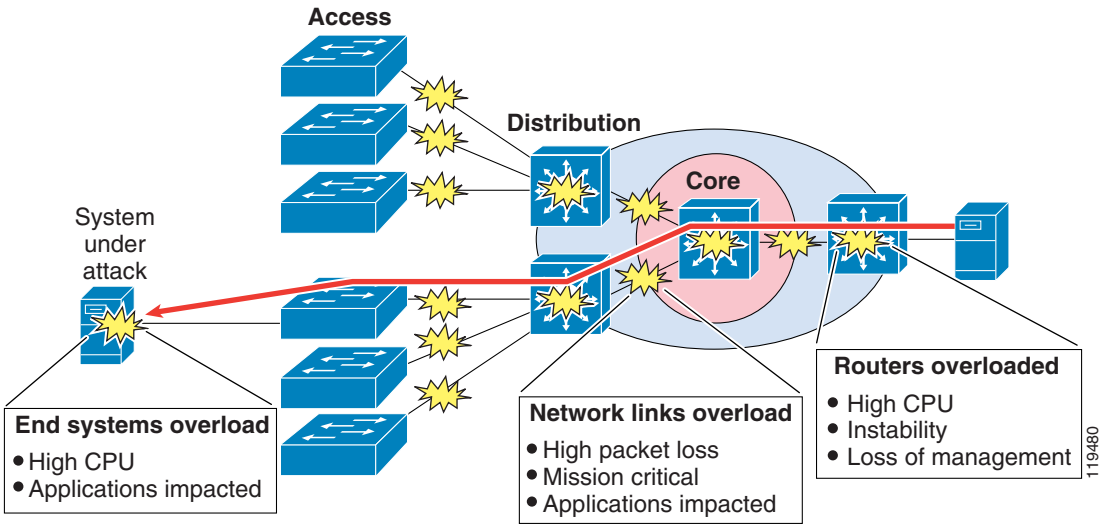
There are two main classes of DoS attacks:

- Spoofing attacks—The attacker pretends to provide a legitimate service, but provides false information to the requester (if any).
- Slamming/flooding attacks—The attacker exponentially generates and propagates traffic until service resources (servers and/or network infrastructure) are overwhelmed.

Spoofing attacks are best addressed by authentication and encryption technologies. Slamming/flooding attacks, on the other hand, can be effectively mitigated through QoS technologies.

Worms, on the other hand, exploit security vulnerabilities in their targets and disguisedly carry harmful payloads that usually include a self-propagating mechanism. Network infrastructure usually isn't 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 CPU/hardware resources of routers and switches in their paths, indirectly causing Denial of Service to legitimate traffic flows, as shown in Figure 1-11.

Figure 1-11 Direct and Indirect Collateral Damage from DoS/Worm Attacks



A *reactive approach* to mitigating such attacks is to reverse-engineer the worm and set up intrusion detection mechanisms and/or ACLs and/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 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.

Note Interestingly, a 2002 CSI/FBI report stated that the majority of network attacks occur from *within* an organization, typically by disgruntled employees. This underscores the need to protect the Access-Edges of enterprise networks as well as their Internet edges.

A *proactive approach* to mitigating DoS/worm flooding attacks within enterprise networks is to immediately respond to out-of-profile network behavior indicative of a DoS or worm attack using Campus Access-Layer policers. Such policers meter traffic rates received from endpoint devices and markdown excess traffic spikes to the Scavenger class (DSCP CS1) when these exceed specified watermarks (at which point they are no longer considered normal flows).

In this respect, the policers are relatively dumb. They do not match specific network characteristics of specific types of attacks, but simply meter traffic volumes and respond to abnormally high volumes as close to the source as possible. The simplicity of this approach negates the need for the policers to be programmed with knowledge of the specific details of how the attack is being generated or propagated.

It is precisely this dumbness of such access layer policers that allow them to maintain relevancy as worms mutate and become more complex. The policers do not care *how* the traffic was generated or *what* it looks like, they care only *how much* traffic is being put onto the wire. Therefore, they continue to police even advanced worms that continually change the tactics of how traffic is being generated.

For example, in most enterprises it is quite abnormal (within a 95 % statistical confidence interval) for PCs to generate sustained traffic in excess of 5 % of link capacity. In the case of a FastEthernet switch port, this means that it would be unusual in most organizations for an end-user PC to generate more than 5 Mbps of *uplink* traffic on a sustained basis.

Note It is important to recognize that this value (5 percent) for normal Access-Edge utilization by endpoints is just an example value. This value would likely vary within the enterprise and from enterprise to enterprise.

It is very important to recognize that what is being proposed is not to police all traffic to 5 Mbps and automatically drop the excess. Should that be the case, there would not be much reason to deploy FastEthernet or GigabitEthernet switch ports to endpoint devices, because even 10-BaseT Ethernet switch ports have more uplink capacity than a 5 Mbps policer-enforced limit. Furthermore, such an approach would supremely penalize legitimate traffic that did happen to exceed 5 Mbps on an FE switch port.

A less draconian approach would be to couple Access Layer policers with hardware/software (Campus/WAN/VPN) queuing polices, with both sets of policies provisioning a “less-than-Best-Effort” Scavenger class. Access Layer policers would markdown out-of-profile traffic to DSCP CS1 (Scavenger) and then have all congestion management policies (whether in Catalyst hardware or in Cisco IOS software) provision a “less-than Best-Effort” queueing service for any traffic marked to DSCP CS1.

Let’s illustrate how this might work for both *legitimate* traffic exceeding the Access Layer’s policer watermark and also the case of *illegitimate* excess traffic resulting from a DoS or worm attack.

In the former case, assume that the PC generates over 5 Mbps of traffic, perhaps because of a large file transfer or backup. Congestion (under normal operating conditions) is rarely if ever experienced within the Campus because there is generally abundant capacity to carry the traffic. Uplinks to the Distribution and Core layers of the Campus network are typically GigabitEthernet and would require 1000 Mbps of traffic from the Access Layer switch to congest. If the traffic is destined to the far side of a WAN/VPN link (which is rarely over 5 Mbps in speed), dropping occurs even without the Access Layer policer, because of the bottleneck caused by the Campus/WAN speed mismatch. The TCP sliding windows mechanism would eventually find an optimal speed (under 5 Mbps) for the file transfer. Access Layer policers that markdown out-of-profile traffic to Scavenger (CS1) would thus not affect legitimate traffic, aside from the obvious remarking. *No reordering or dropping would occur on such flows as a result of these policers that would not have occurred anyway.*

In the latter case, the effect of Access Layer policers on traffic caused by DoS or worm attacks is quite different. As hosts become infected and traffic volumes multiply, congestion may be experienced even within the Campus. If just 11 end-user PCs on a single switch begin spawning worm flows to their maximum FastEthernet link capacities, the GigabitEthernet uplink from the Access Layer switch to the Distribution Layer switch will congest and queuing/reordering/dropping will engage. At this point, VoIP, critical data applications, and even Best Effort applications would gain priority over worm-generated traffic (as Scavenger traffic would be dropped the most aggressively). Furthermore,

network devices would remain accessible for administration of the patches/plugs/ACLs/NBAR-policies required to fully neutralize the specific attack. WAN/VPN links would also be protected: VoIP, critical data and even Best Effort flows would receive priority over any WAN/VPN traffic marked down to Scavenger/CS1. This is a huge advantage, because generally WAN/VPN links are the first to be overwhelmed by DoS/worm attacks. Scavenger-class Access Layer policers thus *significantly mitigate network traffic generated by DoS or worm attacks.*

It is important to recognize the distinction between mitigating an attack and preventing it entirely. The strategy described in this document ***does not guarantee that no DoS or worm attacks will ever happen, but serves only to reduce the risk and impact that such attacks could have on the network infrastructure***.

Scavenger-class QoS DoS/Worm Mitigation Strategy

Let's recap the most important elements of the Scavenger-class QoS DoS/worm mitigation strategy.

First, network administrators need to ***profile applications to determine what constitutes normal as opposed to abnormal flows, within a 95 percent confidence interval***. Thresholds demarking normal/abnormal flows will 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 presented Scavenger- class QoS strategy. Only sustained, abnormal streams generated simultaneously by multiple hosts (highly-indicative of DoS/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 (DSCP CS1)***. Additionally, whenever Cisco Catalyst 6500s with Supervisor 720s are deployed in the distribution layer, ***deploy a second line of policing-defense at the distribution layer via Per-User Microflow Policing***.

To complement these remarking policies, it is necessary to ***enforce end-to-end "less-than-Best-Effort" Scavenger-class queuing policies*** within the Campus, WAN and VPN.

It is critically important to recognize, that even when Scavenger class QoS has been deployed end-to-end, this strategy only *mitigates* DoS/worm attacks, and *does not prevent them or remove them entirely.* Therefore, it is vital to overlay security, firewall, intrusion detection, identity, Cisco Guard, Cisco Traffic Anomaly Detector and Cisco Security Agent solutions in addition to QoS-enabled infrastructures.

Summary

This document began by reviewing **Why Quality of Service is important to Enterprise networks**, specifically because as it enables the transparent convergence of voice, video and data onto a single network. Furthermore, this proven technology can be used to mitigate the impact of DoS/worm attacks.

To set a context for discussion, the QoS toolset was reviewed, including classification and marking tools, policing tools, scheduling tools, link specific tools and AutoQoS.

The next section examined **How is QoS is optimally deployed within the enterprise?** The answer consisted of a 5 phase approach:

1) Strategically defining the objectives to be achieved via QoS—Successful QoS deployments begin by clearly defining organizational QoS objectives and then selecting an appropriate number of service classes to meet these objectives. This section introduced Cisco's QoS Baseline as a strategic guide for selecting the number and type of traffic classes to meet organizational objectives; also a migration

strategy was presented to illustrate how enterprises could start with simple QoS models and gradually increase their complexity as future needs arose. Executive endorsement is recommended, especially when choosing the select few applications to be serviced by the Mission-Critical data class.

2) Analyzing the service-level requirements of the various traffic classes to be provisioned for—the service level needs of voice, video, data and the control plane were discussed. Some of these highlights include the following:

Voice requires 150 ms one-way, end-to-end (mouth-to-ear) delay, 30 ms of one-way jitter and no more than 1 % packet loss. Voice should receive strict priority servicing, and the amount of priority bandwidth assigned for it should take into account the VoIP codec, the packetization rate, IP/UDP/RTP headers (compressed or not) and Layer 2 overhead. Additionally, provisioning QoS for IP telephony requires that a minimal amount of guaranteed bandwidth be allocated to Call-Signaling traffic.

Video comes in two flavors: Interactive Video and Streaming Video. Interactive Video has the same service level requirements as VoIP because a voice call is embedded within the video stream. Streaming Video has much laxer requirements, because of the high amount of buffering that has been built into the applications.

Control plane requirements, such as provisioning moderate bandwidth guarantees for IP Routing and Network Management protocols, should not be overlooked.

Data comes in a variety of forms, but can generally be classified into four main classes: Best Effort (the default class), Bulk (non-interactive, background flows), Transactional/Interactive (interactive, foreground flows) and Mission-Critical. Mission-Critical Data applications are locally-defined, meaning that each organization must determine the select few Transactional Data applications that contribute the most significantly to their overall business objectives.

3) Designing and testing QoS policies prior to production-network deployment—several best-practice QoS design principles were presented to help simplify and streamline the QoS design phase. These included: always performing QoS policies in hardware – rather than software – whenever a choice exists, classifying and marking (with standards-based DSCP markings) as close to the source as technically and administratively feasible, policing as close to the source as possible, and queuing on every node that has a potential for congestion. Queuing guidelines also included not provisioning more than 33% of a link for realtime traffic and reserving at least 25% of a link for the default Best Effort class.

4) Rolling-out the tested QoS designs to the production-network – Once the QoS designs have been finalized and PoC tested, it is vital ensure that the networking team thoroughly understands the QoS features and syntax before enabling features on production networks. Furthermore, it is recommended to schedule network downtime in order to rollout QoS features. A pilot network-segment can be selected for an initial deployment, and pending observation, the rollout can be expanded in stages to encompass the entire enterprise. A rollback strategy is always recommended, to address unexpected issues arising from the QoS deployment.

5) Monitoring service levels to ensure that the QoS objectives are being met – Implementing a QoS solution is not a one-time task that is complete upon policy deployment. A successful QoS policy rollout is followed by ongoing monitoring of service levels and periodic adjustments and tuning of QoS policies.

As business conditions change, the enterprise may need to adapt to these changes and may be required to begin the QoS deployment cycle anew, by redefining their objectives, tuning and testing corresponding designs, rolling these new designs out and monitoring them to see if they match the redefined objectives.

The document concluded by addressing the highly-relevant question: **How can QoS tools be used to mitigate DoS/Worm Attacks?**

A “less-than-Best-Effort” traffic class, called Scavenger, was introduced, and a strategy for using this class for DoS/worm mitigation was presented. Specifically, flows can be monitored and policed at the Campus Access-Edge (and also at the Distribution Layer if Catalyst 6500s with Supervisor 720s are used). Out-of-profile flows can be marked down to the Scavenger marking (of DSCP CS1). To complement these policers, queues providing a “less-than-Best-Effort” Scavenger service during periods of congestion can be deployed in the LAN, WAN and VPN. Such a strategy would not penalize legitimate traffic flows that were temporarily out of profile; however sustained abnormal streams, highly-indicative of DoS/worm attacks, would be subject to aggressive dropping only after legitimate traffic was fully serviced.

It is critically important to recognize, that even when Scavenger-class QoS has been deployed end-to-end, this strategy only mitigates DoS/worm attacks, and does not prevent them or remove them entirely. Therefore, it is vital to overlay security, firewall, intrusion detection, identity, Cisco Guard, Cisco Traffic Anomaly Detctor and Cisco Security Agent solutions in addition to QoS-enabled infrastructures.

References

Standards

- RFC 791 “Internet Protocol Protocol Specification” http://www.ietf.org/rfc/rfc791
- RFC 2474 "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers http://www.ietf.org/rfc/rfc2474
- RFC 2597 "Assured Forwarding PHB Group" http://www.ietf.org/rfc/rfc2597
- RFC 2697 "A Single Rate Three Color Marker" http://www.ietf.org/rfc/rfc2697
- RFC 2698 "A Two Rate Three Color Marker" http://www.ietf.org/rfc/rfc2698
- RFC 3168 "The Addition of Explicit Congestion Notification (ECN) to IP" http://www.ietf.org/rfc/rfc3168
- RFC 3246 "An Expedited Forwarding PHB (Per-Hop Behavior)" http://www.ietf.org/rfc/rfc3246

Books

- Szigeti, Tim and Christina Hattingh. *End-to-End QoS Network Design: Quality of Service in LANs, WANs and VPNs*. Indianapolis: Cisco Press, 2004.

Cisco Documentation

- Cisco IOS QoS Configuration Guide – Cisco IOS version 12.3 http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123cgcr/qos_vcg.htm
- Cisco IOS Configuration Guide – Configuring Data Link Switching Plus - Cisco IOS version 12.3 http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fibm_c/bcfpart2/bcfdlsw.htm
- Understanding How Routing Updates and Layer 2 Control Packets Are Queued on an Interface with a QoS Service Policy (PAK_PRIORITY) http://www.cisco.com/warp/public/105/rtgupdates.html

CHAPTER 2

Campus QoS Design

This chapter includes the following topics:

QoS Design Overview

This section includes the following topics:

- Where is QoS Needed in a Campus?
- DoS/Worm Mitigation
- Call Signaling Ports
- Access Edge Trust Models
- Cisco IP Phones
- WAN Aggregator/Branch Router Connections

Where is QoS Needed in a Campus?

The case for Quality of Service (QoS) in WANs/VPNs is largely self-evident because of its low-bandwidth links compared to the high-bandwidth requirements of most applications. However, the need for QoS is sometimes overlooked or even challenged in high-bandwidth Gigabit/TenGigabit campus LAN environments.

Although network administrators sometimes equate QoS only with queuing, the QoS toolset extends considerably beyond just 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 (access edge).

Three important 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/Per-Hop Behaviors. Sometimes endpoints can be trusted to set Class of Service (CoS)/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, a user 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 (like VoIP) throughout the enterprise. For this reason, the clause "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.

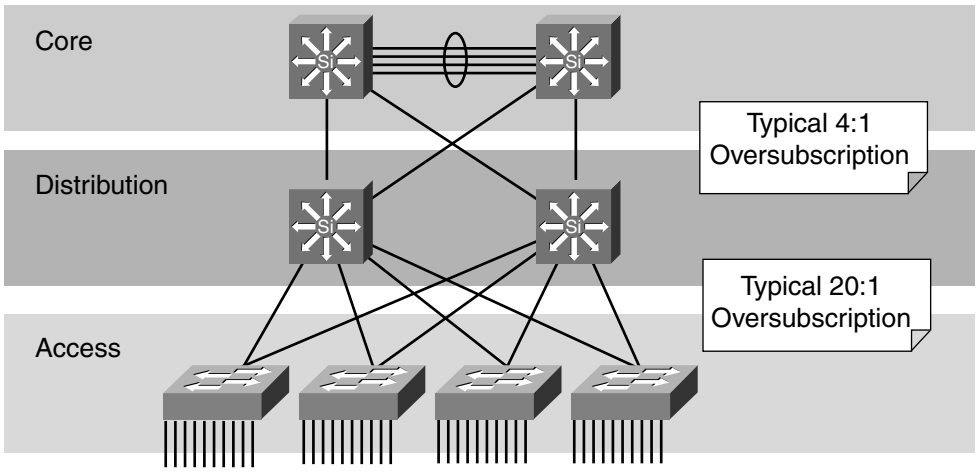
  Cisco IOS routers perform QoS in software. This places additional demands on the CPU, depending on the complexity and functionality of the policy. Cisco Catalyst switches, on the other hand, perform QoS in dedicated hardware ASICS and as such do not tax their main CPUs to administer QoS policies. You can therefore apply complex QoS policies at Gigabit/TenGigabitEthernet 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, 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, as shown in Figure 2-1.

Figure 2-1 Typical Campus Oversubscription Ratios



It is quite rare under normal operating conditions for campus networks to suffer congestion. And 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, in fact, this may occur. The potential for congestion exists in campus uplinks because of oversubscription ratios and speed mismatches in campus downlinks (for example, GigabitEthernet to FastEthernet 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/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, as detailed later in this chapter, maintains network availability by protecting and servicing critical applications such as VoIP and even Best Effort traffic.

The intrinsic interdependencies of network QoS, High Availability and security are clearly manifest in such worse-case scenarios.

So where is QoS required in campus?

Access switches require the following QoS policies:

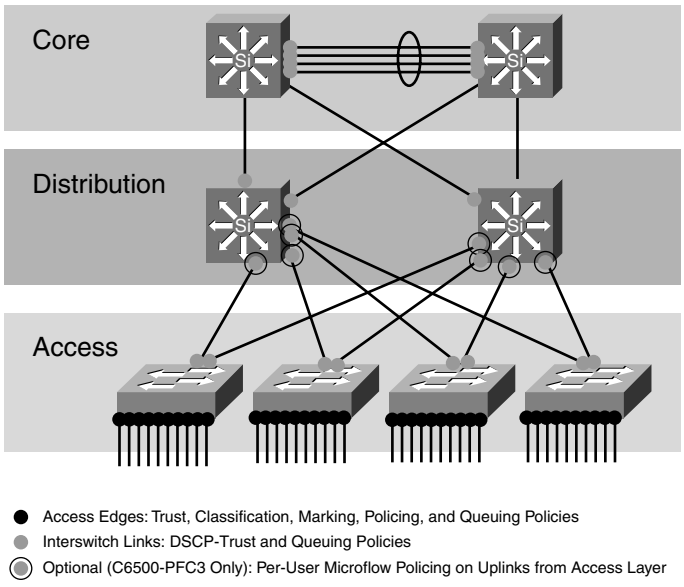
- Appropriate (endpoint-dependant) trust policies, and/or classification and marking policies
- Policing and markdown policies
- Queuing policies.

Distribution and core switches require the following QoS policies:

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

These recommendations are summarized in Figure 2-2.

Figure 2-2 Where is QoS Required within the Campus?



DoS/Worm Mitigation Strategies

A proactive approach to mitigating DoS/worm flooding attacks within campus environments is to immediately respond to out-of-profile network behavior indicative of a DoS or worm attack through access layer policers. Such policers meter traffic rates received from endpoint devices and markdown excess traffic when these exceed specified watermarks (at which point they are no longer considered normal flows).

These policers are relatively "dumb" because they do not match specific network characteristics of specific types of attacks. Instead, they simply meter traffic volumes and respond to abnormally high volumes as close to the source as possible. The simplicity of this approach negates the need for the policers to be programmed with knowledge of the specific details of *how* the attack is being generated or propagated.

It is precisely this "dumbness" of such access layer policers that allow them to stay effective as worms mutate and become more complex. The policers do not care *how* the traffic was generated or *what* it looks like, they only care *how much* traffic is being put onto the wire. Therefore, they continue to police even advanced worms that continually change the tactics of how traffic is being generated.

For example, in most enterprises it is quite abnormal (within a 95 percent statistical confidence interval) for PCs to generate sustained traffic in excess of 5 percent of their link capacity. In the case of a FastEthernet switch port, this means that it is unusual in most organizations for an end-user PC to generate more than 5 Mbps of uplink traffic on a sustained basis.

Note It is important to recognize that this value for normal access edge utilization by endpoints of 5 percent is just an *example* value used for simplicity in this chapter. This value would likely vary from industry vertical to vertical, and from enterprise to enterprise.

Cisco does not recommend policing all traffic to 5 Mbps and automatically dropping the excess. If that was the case, there would be no need to deploy FastEthernet or GigabitEthernet switch ports to endpoint devices because even 10-BaseT Ethernet switch ports have more uplink capacity than a 5 Mbps policer-enforced limit. Such an approach would also penalize legitimate traffic that *did* exceed 5 Mbps on a FastEthernet switch port.

A less severe approach is to couple access layer policers with hardware/software (campus/WAN/VPN) queuing polices, with both sets of policies provisioning for a less-than Best-Effort Scavenger class.

With this method, access layer policers mark down out-of-profile traffic to DSCP CS1 (Scavenger) and then all congestion management policies (whether in Catalyst hardware or in IOS software) provision a less-than Best-Effort service for any traffic marked to CS1 during periods of congestion.

This method works for both legitimate traffic exceeding the access layer policer's watermark and also for illegitimate excess traffic as a result of a DoS or worm attack.

In the former case, for example, assume that the PC generates over 5 Mbps of traffic, perhaps because of a large file transfer or backup. Congestion under normal operating conditions is rarely if ever experienced because there is generally abundant capacity to carry the traffic within the campus. This is usually true because the uplinks to the distribution and core layers of the campus network are typically GigabitEthernet and require 1000 Mbps of traffic from the access layer switch to congest. If the traffic is destined to the far side of a WAN/VPN link, which is rarely over 5 Mbps in speed, dropping occurs even without the access layer policer, because of the campus/WAN speed mismatch and resulting bottleneck. TCP's sliding windows mechanism eventually finds an optimal speed (under 5 Mbps) for the file transfer. Thus, access layer policers that mark down out-of-profile traffic to Scavenger (CS1) *do not affect legitimate traffic*, aside from the obvious remarking. No reordering or dropping occurs on such flows as a result of these policers that would not have occurred in any event.

In the case of illegitimate excess traffic, the effect of access layer policers on traffic caused by DoS or worm attacks is quite different. As many hosts become infected and traffic volumes multiply, congestion may be experienced in the campus up-links due to the aggregate traffic volume. For example, if just 11 end-user PCs on a single access-layer switch begin spawning worm flows to their maximum FastEthernet link capacities, the GigabitEthernet up-link to the distribution layer switch will congest, and queuing/reordering will engage. At such a point, VoIP and critical data applications, and even Best Effort applications, gain priority over worm-generated traffic. Scavenger traffic is dropped the most aggressively, while network devices remain accessible for the administration of patches/plugs/ACLs required to fully neutralize the specific attack.

WAN links are also protected. VoIP, critical data and even Best Effort flows continue to receive priority over any traffic marked down to Scavenger/CS1. This is a huge advantage, because WAN links are generally the first to be overwhelmed by DoS/worm attacks. Access layer policers thus *significantly mitigate* network traffic generated by DoS or worm attacks.

You should recognize the distinction between mitigating an attack and preventing it entirely. The strategy presented here *does not guarantee that no Denial of Service or worm attacks will ever happen, but serves only to reduce the risk and impact* that such attacks have on the campus network infrastructure and then, by extension, the WAN/VPN network infrastructure. Furthermore, while this strategy reduces the collateral damage to the network infrastructure caused by DoS/worm attacks, it may not mitigate other specific objectives of such worms, such as reconnaissance and vulnerability exploitation. Hence, a comprehensive approach much be used to address DoS/worm attacks, involving a holistic integration of security technologies with Quality of Service technologies.

Call Signaling Ports

In this design chapter, only Skinny Call Control Protocol (SCCP) ports (TCP Ports 2000–2002) are used to identify call signaling protocols to keep the examples relatively simple.

However, SCCP is by no means the only call signaling protocol used in IP telephony environments. Cisco recommends including all relevant call signaling ports required for a given IPT environment in the access lists that identify call signaling protocols. Firewalls protecting CallManagers should also allow additional ports to provide the supplementary services that CallManagers provide or require.

Access Edge Trust Models

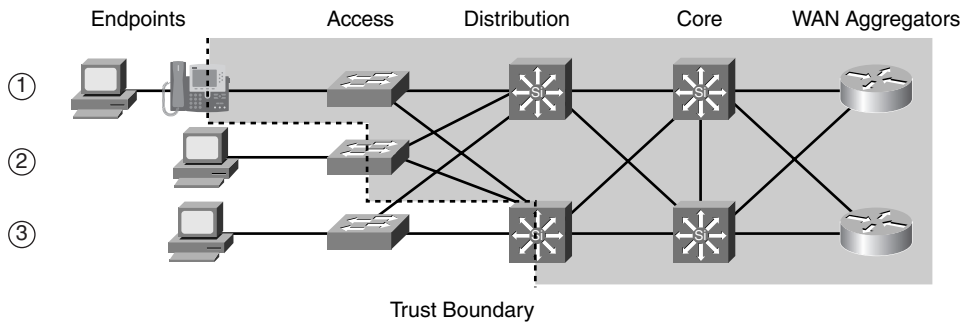
This section includes the following topics:

- Trusted Endpoints
- Untrusted Endpoints
- 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 Figure 2-3.

Figure 2-3 Establishing Trust Boundaries



Legend for Figure 2-3:

1. Optimal trust boundary: trusted endpoint
2. Optimal trust boundary: untrusted endpoint
3. Sub-optimal trust boundary

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

Trusted endpoints have the capabilities and intelligence to mark application traffic to the appropriate CoS and/or DSCP values. Trusted endpoints also have the ability to remark 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.

  Examples of analog gateways include the following:

  - Analog network modules (NM-1V and NM2-V, which support either high- or low-density Voice/Fax Interface Cards (VICs)
  - Cisco Communication Media Module (CMM) linecard
  - Catalyst 6500 Analog Interface Module (WS-X6624-FXS).
  - Cisco VG224 and VG248 IOS-based voice 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 7935 and 7936.
- Videoconferencing gateways and systems—These devices transmit interactive video across the IP network. Examples of such devices capable of setting DSCP markings include the Cisco IP/VC 3511, 3521, 3526 and 3540 videoconferencing gateways and systems. If, on the other hand, video-conferencing devices do not have the ability to set DSCP markings correctly, they should be treated as untrusted devices.
- 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. These 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, the network administrator 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, at best, and usually involves considerable finger pointing. Additionally, network administrators should bear in mind that the majority of DoS/ worm attacks target servers. Infected servers not only might spew profuse amounts of traffic onto the network, but, in such cases, they might do so with trusted markings. There's no hard-and-fast rule that will apply to every situation. Some administrators prefer to trust certain servers, like Cisco CallManagers, due to the large number of ports that may be in use to provide services rather than administer complex access lists. In either case, consider the tradeoffs involved when deciding whether or not to trust a server.
- Wireless access points—Some wireless access points (APs) have the ability to mark or remark 802.1p CoS and/or 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. Examples include the Cisco 7920G wireless IP Phone.

When trusted endpoints are connected to a switch port, all that is typically required is enabling the following interface command: **mls qos trust dscp**.

Optionally, if the traffic rate of the trusted application is known, the network administrator could 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 kbps 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 kbps, 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.

Untrusted Endpoints

This section includes the following topics:

- Untrusted PC + SoftPhone with Scavenger-Class QoS
- Untrusted Server with Scavenger-Class QoS

As previously mentioned, trusting end users and their PCs is generally a bad idea because newer operating systems like 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 a nightmare. Also, marking application traffic on server NICs has disadvantages as discussed in the previous section that may make it preferable to treat these 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, and so will be examined individually.

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 Cisco IP Softphone. In such a case, the critical application needs to be identified using access lists and marked/remarked at the access edge. Remarking can be done with either an MLS QoS **set ip dscp** command or with a policer.

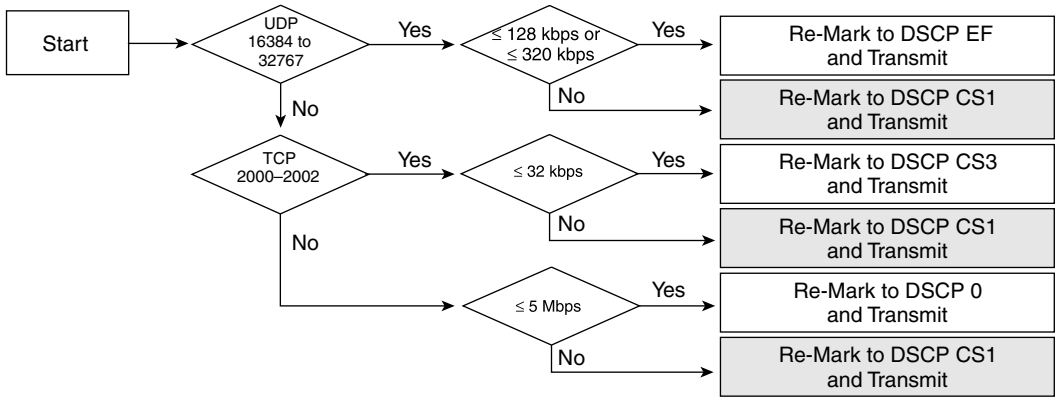
A policer is recommended in this case because limits on the amount of traffic being marked can then be imposed to prevent abuse. Cisco SoftPhones can use regular G.711 codecs, in which case 128 kbps is adequate, or they can be configured use a G.722 (wide codec), in which case 320 kbps is required. The 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 picking random ports within the UDP range of 16383–32767. This is recommended because this allows for a more granular access list to match legitimate Cisco SoftPhone traffic, thereby tightening the overall security of the policy.

Note In this context, "Softphone" can be used to 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 Figure 2-4.

Figure 2-4 Untrusted Endpoint Policing —PC + SoftPhone + Scavenger Model



The syntax for implementing such a policer may vary slightly from platform to platform, as is detailed in the subsequent platform-specific sections.

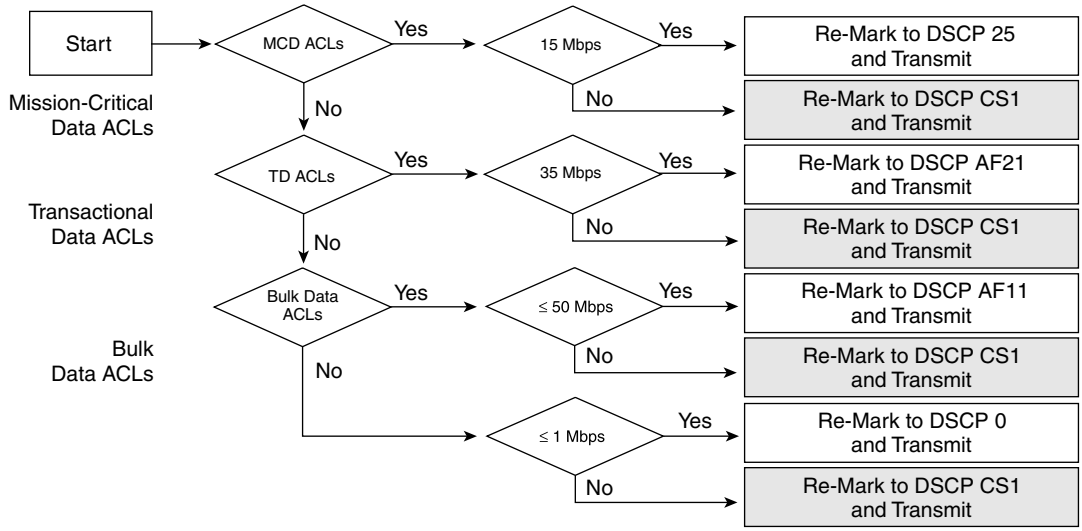
Untrusted Server with Scavenger-Class QoS

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

For an example, assume a single server is running multiple applications, in this case SAP (TCP ports 3200–3203 and also 3600), Lotus Notes (TCP port 1352), and 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 Mbps, 35 Mbps and 50 Mbps, 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. The logic of such a policer is shown in Figure 2-5.

Figure 2-5 Untrusted Endpoint Policing—Multi-Application Server + Scavenger Model



Remember that when deploying QoS designs for untrusted servers, the applications are usually identified by source ports, and not destination ports (as is the case with client-to-server access lists). Thus the access list becomes:

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

as opposed to:

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

This is a subtle but critical difference.

Conditionally-Trusted Endpoints

This section includes the following topics:

- Cisco IP Phones
- Cisco AutoQoS—VoIP
- Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model
- Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model

One of the main business advantages of IP telephony is the simplicity and related cost savings of user adds/moves/changes. To move, a user simply picks up their IP phone, plugs it in at his or her new location and carries 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 get 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).

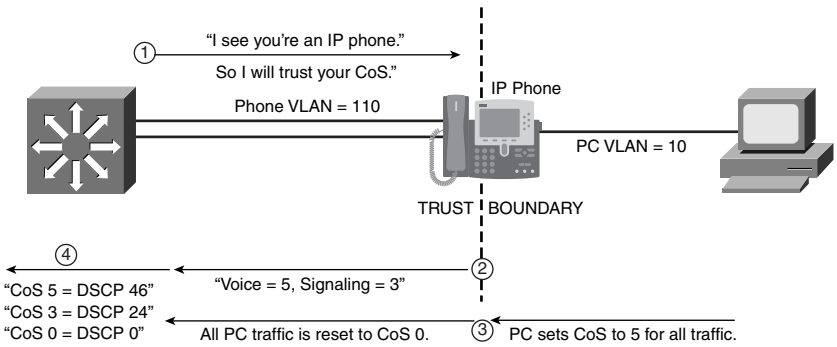
One solution is to place a call to the networking help desk when the move is scheduled, so that the switch ports can be reconfigured to trust/untrust the endpoints as required. However, this approach dampens the mobility business advantage of IP telephony, since manual network administration is then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

Another solution is to have an intelligent exchange of information between the switch and the devices plugged into their ports. If the switch discovers a device that is trustworthy, then it can extend trust to it dynamically; if not, then not.

Cisco IP Phones use the latter solution. In the current Cisco implementation, the intelligent exchange of information is performed using Cisco Discovery Protocol (CDP).

Figure 2-6 shows a conditional trust boundary extension granted to an IP Phone that has passed a CDP exchange.

Figure 2-6 Conditionally-Trusted Endpoint—Trust Boundary Extension and Operation



The sequence shown in Figure 2-6 is the following:

1. Switch and phone exchange CDP; trust boundary is extended to IP Phone.
2. Phone sets CoS to 5 for VoIP and to 3 for call signaling traffic.
3. Phone rewrites CoS from PC to 0.
4. Switch trusts CoS from phone and maps CoS to DSCP for output queuing.

CDP is 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.

Cisco IP Phones

The following overview of some of the main IP Phones helps to explain their impact on access edge QoS design.

- Cisco 7902G— The 7902G is an entry-level IP phone that addresses the voice¬communication needs of areas where only a minimal amount of features is required, such as lobbies, hallways, and break rooms. These phones probably would not be moved. The 7902G has only a single 10BASE-T Ethernet port on the back of the phone; therefore, there is no hardware support to connect a PC to it.

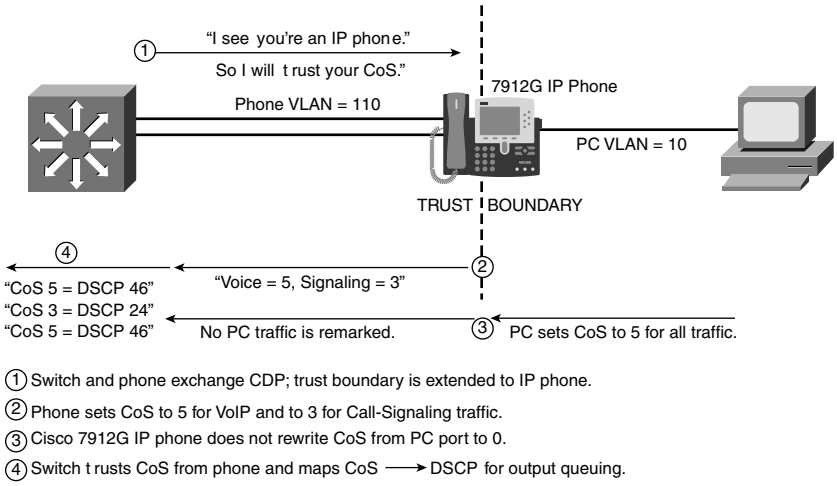
- Cisco 7905G—The 7905G is a basic IP phone that addresses the voice communication needs of a cubicle worker who conducts low to medium telephone traffic. The Cisco 7905G has only a single 10BASE-T Ethernet port on the back of the phone; therefore, there is no hardware support to connect a PC to it.
- Cisco 7910G and 7910G+SW—The 7910G and 7910G+SW IP phones address the voice-communication needs associated with a reception area, lab, manufacturing floor, or employee with a minimal amount of telephone traffic. The only difference between the Cisco 7910G and the Cisco 7910G+SW is that the former has a single 10BASE-T Ethernet port (therefore, there is no hardware support to connect a PC to it), and the latter has two 10/100BASE-T Ethernet ports, which allow a PC to be connected to the IP phone.
- Cisco 7912G—The 7912G is a basic IP phone that addresses the voice-communication needs of a cubicle worker who conducts low to medium telephone traffic. The 7912G supports inline power and an integrated 10/100 Ethernet switch for connecting a PC. The switch used in the 7912G has the capability to mark CoS and DSCP of Voice and Call Signaling traffic that originates from the IP phone, but ***the Cisco 7912G does not have the capability to re-mark CoS values of PC-generated traffic***.
- Cisco 7940G—The 7940G IP phone is suited best for an employee in a basic office cubicle environment—a transaction-type worker, for example—who conducts a medium amount of business by telephone. The 7940G supports inline power and has an integrated 10/100 Ethernet switch for connecting a PC.
- Cisco 7960G—The 7960G is designed to meet the communication needs of a profes¬sional worker in an enclosed office environment—an employee who experiences a high amount of phone traffic in the course of a business day. The 7960G supports inline power and has an integrated 10/100 Ethernet switch for connecting a PC.
- Cisco 7970G—The 7970G not only addresses the needs of the executive or major decision maker, but also brings network data and applications to users without PCs. This IP phone includes a backlit, high-resolution color touch-screen display. Cur¬rently, Cisco 7970G is the only Cisco IP phone that supports both Cisco prestandard Power over Ethernet (PoE) and the IEEE 802.3af PoE. The 7970G has an integrated 10/100 Ethernet switch for connecting a PC.

All of the IP Phones listed above have the ability to mark 802.1Q/p CoS values for both VoIP and call signaling (default values are 5 and 3, respectively). Furthermore, they also have the ability to mark DSCP values for both VoIP and call signaling (current defaults are EF and AF31, respectively; future software releases will change these values to EF and CS3, respectively).

IP Phone models 7902G, 7905G and 7910G lack the hardware to connecting a PC behind the Cisco IP Phone. All other IP Phone models listed above (except the 7912G) have the hardware support to connect a PC behind the IP Phone and also support 802.1Q/p CoS remarking of tagged packets that may originate from such PCs.

The 10/100 Ethernet switch built into the 7912G does not have the support to re-mark CoS values that might have been set by a PC, as illustrated in Figure 2-7. This re-marking limitation represents a potential security hole for enterprises deploying these IP phones. However, this hole can be plugged, for the most part, with access-edge policers, as will be detailed in this chapter. It is important to note that if 7912G IP phones are deployed to users that move locations, all user switch ports within the enterprise should have access-edge policers set on them to ensure mobility and security if a 7912G user moves the phones to another port.

Figure 2-7 Conditionally-Trusted Endpoint—Cisco 7912G Trust Boundary Extension and Operation



The sequence as shown in Figure 2-7 is as follows:

1. Switch and phone exchange CDP; trust boundary is extended to IP Phone
2. Phone sets CoS to 5 for VoIP and to 3 for call signaling traffic.
3. Cisco 7912G IP Phone does not rewrite CoS from PC port to 0
4. Switch trusts CoS from phone and maps CoS to DSCP for output queuing

AutoQoS—VoIP

When the main business objective of the QoS deployment is to enable QoS for IP Telephony only (i.e., without Scavenger-class QoS), then the network administrator may choose to take advantage of the Cisco AutoQoS VoIP feature.

AutoQoS VoIP is essentially an intelligent macro that enables an administrator to enter one or two simple AutoQoS commands to enable all the appropriate features for the recommended QoS settings for an VoIP and IP Telephony for a specific platform and/or a specific interface.

AutoQoS VoIP automatically configures the best-practice QoS configurations (based on previous Cisco Enterprise QoS SRNDs) for VoIP on Cisco Catalyst switches and IOS routers. By entering one global and/or one interface command (depending on the platform), the AutoQoS VoIP macro then would expand these commands into the recommended VoIP QoS configurations (complete with all the calculated parameters and settings) for the platform and interface on which the Auto-QoS VoIP macro is applied.

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/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

The standard (interface) configuration commands to enable AutoQoS are: **auto qos voip**. Depending on the platform, AutoQoS VoIP can support the following additional keyword commands:

- **cisco-phone**—When you enter the **auto qos voip cisco-phone** interface configuration command on a port at the edge of a network that is connected to a Cisco IP Phone, the switch enables the conditional-trusted boundary feature. The switch uses the Cisco Discovery Protocol (CDP) to detect the presence or absence of a Cisco IP Phone. When a Cisco IP Phone is detected, the ingress classification on the interface is set to trust the CoS marking of the received packet. When a Cisco IP Phone is absent, the ingress classification is set to not trust the CoS (or DSCP) value of any packet.
- **cisco-softphone**—When you enter the **auto qos voip cisco-softphone** interface configuration command on a port at the edge of the network that is connected to a device running the Cisco SoftPhone, the switch uses policing to decide whether a packet is in or out of profile and to specify the action on the packet. If the packet does not have a DSCP value of 24, 26, or 46 or is out of profile, the switch changes the DSCP value to 0.
- **trust** —When you enter the **auto qos voip trust** interface configuration command on a port connected to the interior of the network, the switch trusts the CoS value in ingress packets (the assumption is that traffic has already been classified by other edge devices).

The AutoQoS commands and optional keywords are shown on a per-platform basis in the platform-specific design sections of this chapter.

Additionally, it should be pointed out that AutoQoS VoIP can also be viewed as a template which may be modified and expanded on to support additional classes of applications. In this manner, the AutoQoS VoIP feature can be used to quickly and accurately deploy 80% (or more) of the desired solution, which then can be manually customized further to tailor to the specific customer requirements.

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

In this model, trust of CoS markings is extended to CDP-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 family supports per-port/per-VLAN policing as a feature. Other platforms have already committed to supporting this feature in the near future. 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. Such examples are provided later in this chapter.

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

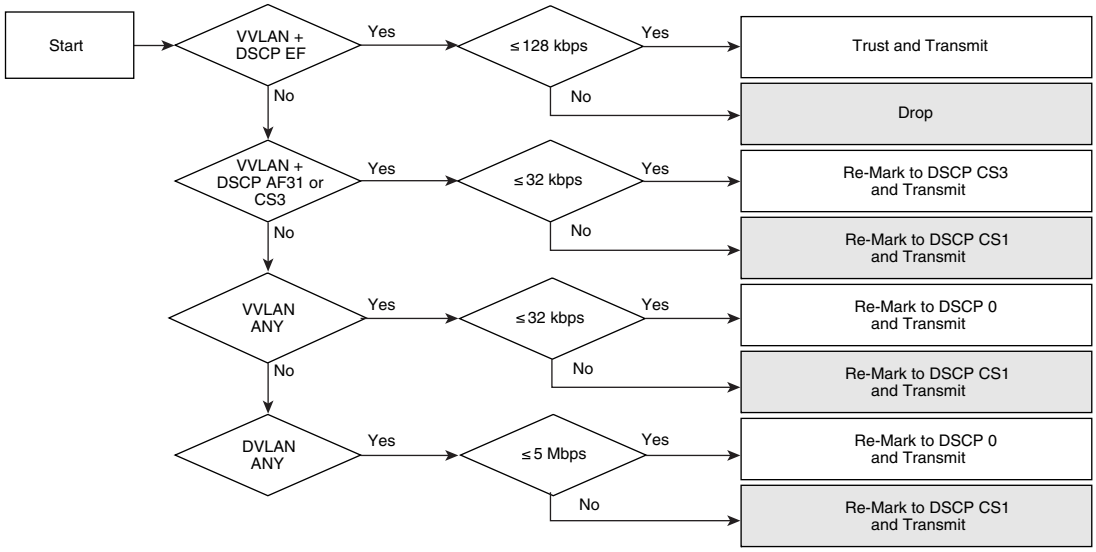
- 128 kbps for Voice traffic, marked CoS 5/DSCP EF (320 kbps in the case of G.722 codecs)
- 32 kbps for call signaling traffic (marked CoS 3/DSCP AF31 or CS3)
- 32 kbps of Best Effort services traffic (marked CoS 0)

There should not be any other traffic originating from the VVLAN, so 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 Mbps for FE ports) to Scavenger/CS1.

The logic of these policers is shown in Figure 2-8.

Figure 2-8 Conditionally-Trusted Endpoint Policing—IP Phone + PC + Scavenger (Basic) Model



Conditionally-Trusted IP Phone + 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 SoftPhone. If the UDP ports used by the desktop videoconferencing application can be explicitly defined within the application, as with SoftPhone, then you can use two policers: one for IP/VC and another for SoftPhone. Otherwise, a single policer covering the UDP port range of 16384–32767 is required, which would be provisioned for the worst-case scenario of legitimate traffic. In this case, this is the videoconferencing application's requirement of 500 kbps (for a 384 kbps desktop IP/VC application), as compared to the SoftPhone requirement of 128 kbps (or 320 kbps for G.722 codecs).

Note Policer thresholds should be set according to the video application's requirements. Some interactive-video applications may have higher bandwidth requirements for their codecs; for example Cisco Video Telephony (VT) Advantage includes a proprietary codec that requires 7 Mbps of bandwidth.

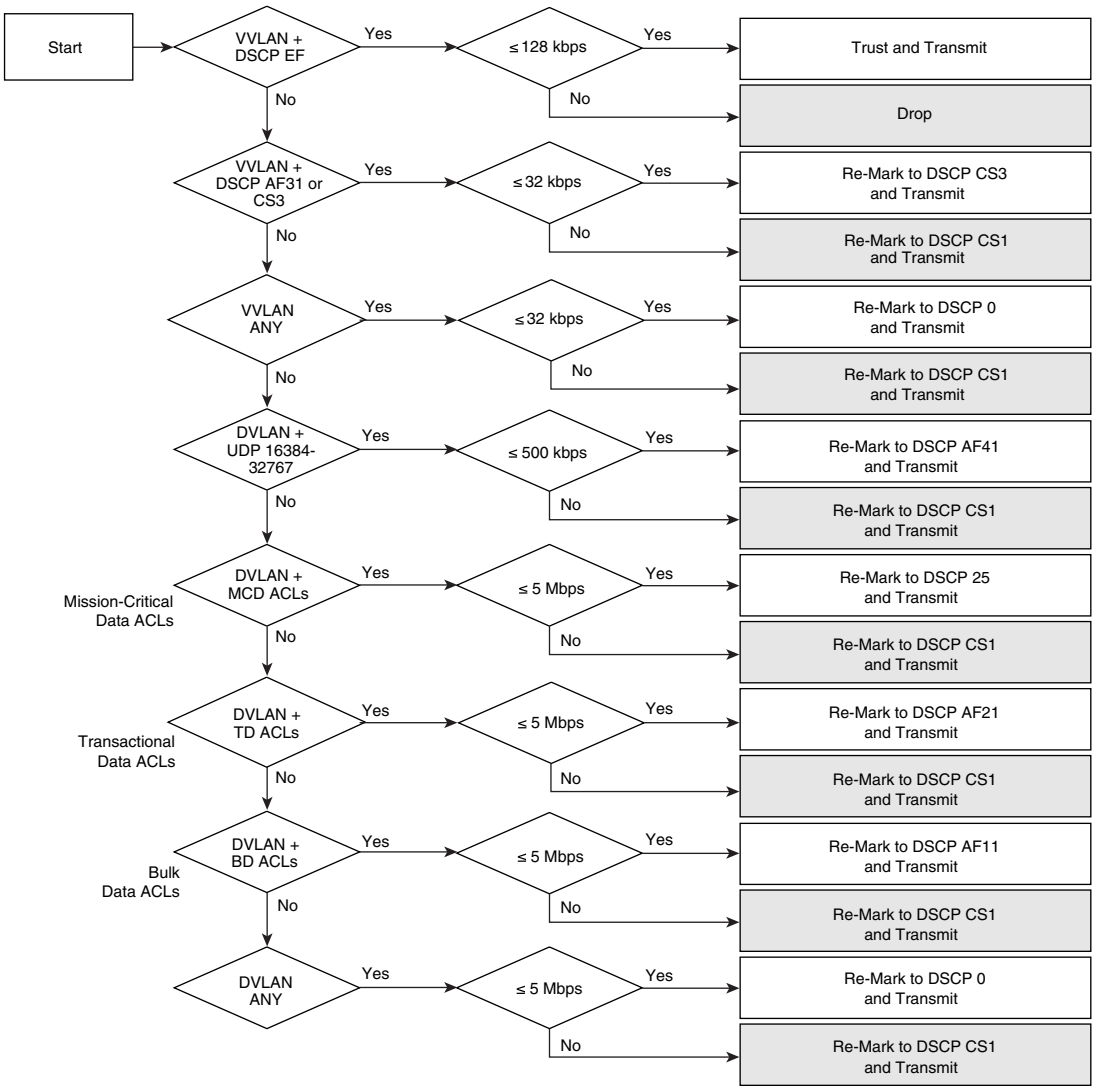
You can add additional data VLAN policers to meter Mission-Critical Data, Transactional Data and Bulk Data flows. Each of these classes can be policed on ingress to the switch port to an in-profile amount, such as 5 percent each.

Note Since Mission-Critical and Transactional Data applications are interactive foreground applications requiring user input, it is highly unlikely that both of these types of applications will be simultaneously generating 5 Mbps each from a client PC. However, in the rare case that they are, these flows will be policed further by any per-user microflow policing policies that may be deployed on distribution layer Catalyst 6500 Supervisor 720s (PFC3s), as is detailed later in this chapter.

Another factor to keep in mind is that certain Catalyst platforms allow only up to 8 policers per FastEthernet 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 access lists, which is detailed in the configuration examples presented later in this chapter.

The logic of these advanced policers is shown in Figure 2-9.

Figure 2-9 Conditionally-Trusted Endpoint Policing—IP Phone + PC + Scavenger (Advanced) Model



Legend for Figure 2-9:

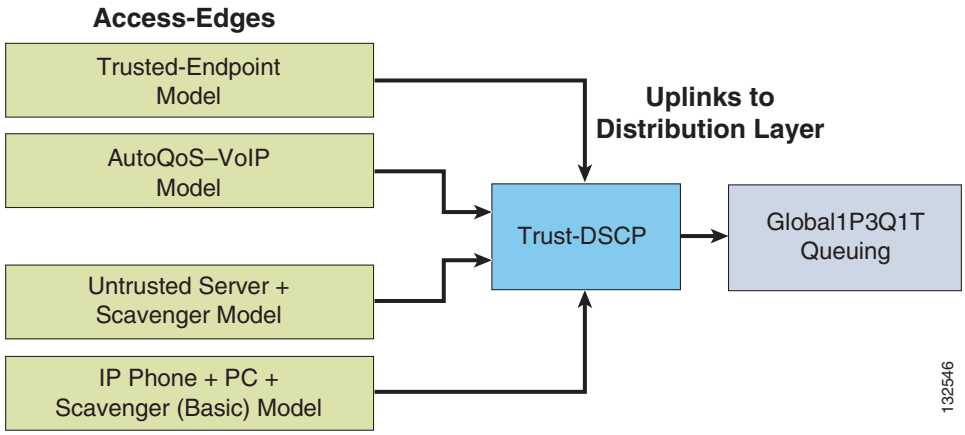
- MCD = Mission Critical Data
- TD = Transactional Data
- BD = Bulk Data

Catalyst 2950—QoS Considerations and Design

This section includes the following topics:

The Catalyst 2950 does not support Layer 3 forwarding, and as such is only applicable as a low-end access layer switch. The QoS design options for a Catalyst 2950 are shown in Figure 2-10.

Figure 2-10 Access Layer Catalyst 2950 QoS Design Options



Cisco recommends using the Enhanced Image (EI) versions of IOS software on these platforms because these offer additional QoS features such as MQC/ACL classification options, policing and markdown functions, mapping tables and AutoQoS.

Catalyst 2950—Trusted Endpoint Model

This section includes the following topics:

Configuration

Configuring a Catalyst 2950 to trust an endpoint is fairly straightforward, as shown below. The trusted endpoint should be assigned either the voice VLAN (VVLAN) or the data VLAN (DVLAN) with the appropriate switchport commands.

Example 2-1 Catalyst 2950—Trusted Endpoint Model Configuration

```
CAT2950(config)#interface FastEthernet0/1
CAT2950(config-if)#mls qos trust dscp
```

Catalyst MLS QoS Verification Command

The **show mls qos interface** verification command reports the configured trust state and the current operating trust mode of a switchport interface.

In this example, the command verifies that interface FastEthernet 0/1 correctly trusts the DSCP values of the endpoint to which it is connected.

Example 2-2 Show MLS QoS Interface Verification of a Switchport Connected to a Trusted Endpoint

```
CAT2950#show mls qos interface FastEthernet0/1
FastEthernet0/1
trust state: trust dscp               ! Configured trust state is to trust DSCP
trust mode: trust dscp                ! Current operating mode is to trust DSCP
COS override: dis
default COS: 0
pass-through: none
trust device: none
CAT2950#
```

Catalyst 2950—AutoQoS VoIP Model

The Catalyst 2950 supports AutoQoS VoIP with the following keyword options:

- **auto qos voip cisco-phone**
- **auto qos voip cisco-softphone**
- **auto qos voip trust**

When you enable AutoQoS VoIP on the Catalyst 2950 by using the **auto qos voip cisco-phone**, the **auto qos voip cisco-softphone**, or the **auto qos voip trust** interface configuration command, the switch automatically generates a QoS configuration based on the traffic type and ingress packet label and applies the commands listed in Table 2-1 to the interface.

Table 2-1 Catalyst 2950 Auto-QoS Generated Configuration

Description	Automatically Generated QoS Command Equivalent
The switch automatically enables standard QoS and configures the CoS-to-DSCP map (maps CoS values in incoming packets to a DSCP value).	`CAT2950(config)# mls qos map cos-dscp 0 8 16 26 32 46 48 56`
If you entered the **auto qos voip trust** command, the switch automatically sets the ingress classification on the interface to trust the CoS value received in the packet.	`CAT2950(config-if)# mls qos trust cos`

Table 2-1 Catalyst 2950 Auto-QoS Generated Configuration

If you entered the **auto qos voip cisco-phone** command, the switch automatically enables the trusted boundary feature, which uses the CDP to detect the presence or absence of a Cisco IP Phone.	`CAT2950(config-if)# mls qos trust device cisco-phone`

If you entered the **auto qos voip cisco-softphone** command, the switch automatically creates class maps and policy maps.	`CAT2950(config)# class-map match-all AutoQoS-VoIP-RTP-Trust`
	`CAT2950(config-cmap)# match ip dscp 46`
	`CAT2950(config)# class-map match-all AutoQoS-VoIP-Control-Trust`
	`CAT2950(config-cmap)# match ip dscp 24 26`
	`CAT2950(config)# policy-map AutoQoS-Police-SoftPhone`
	`CAT2950(config-pmap)# class AutoQoS-VoIP-RTP-Trust`
	`CAT2950(config-pmap-c)# set ip dscp 46`
	`CAT2950(config-pmap-c)# police 1000000 4096 exceed-action drop`
	`CAT2950(config-pmap)# class AutoQoS-VoIP-Control-Trust`
	`CAT2950(config-pmap-c)# set ip dscp 24`
	`CAT2950(config-pmap-c)# police 1000000 4096 exceed-action drop`

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.	`CAT2950(config-if)# service-policy input AutoQoS-Police-SoftPhone`

The switch automatically assigns egress queue usage on this interface.
The switch enables the egress expedite queue and assigns WRR weights to queues 1, 2, and 3. (The lowest value for a WRR queue is 1. When the WRR weight of a queue is set to 0, this queue becomes an expedite queue.)
The switch configures the CoS-to-egress-queue map:
• CoS values 0 and 1 select queue 1.
• CoS values 2 and 4 select queue 2.
• CoS values 3, 6, and 7 select queue 3.
• CoS value 5 selects queue 4.
	`CAT2950(config)# wrr-queue bandwidth 10 20 70 1`
	`CAT2950(config)# no wrr-queue cos-map`
	`CAT2950(config)# wrr-queue cos-map 1 0 1`
	`CAT2950(config)# wrr-queue cos-map 2 2 4`
	`CAT2950(config)# wrr-queue cos-map 3 3 6 7`
	`CAT2950(config)# wrr-queue cos-map 4 5`

Catalyst 2950—Untrusted PC + SoftPhone with Scavenger-Class QoS Model

The Catalyst 2950 does not support the **range** keyword within an ACL when the ACL is being referenced by a MQC class-map. Therefore, a policy to mark UDP flows in the port range of 16384 through 32767 cannot be configured on the Catalyst 2950.

A possible workaround to this limitation would be to pre-set the port(s) to be used by SoftPhone within the application itself. In such a case, these ports would have to be discretely matched by ACL entries on the Catalyst 2950. Furthermore, each port being used for call signaling would also require a discrete ACL entry.

However, even in the case where all these ports are buttoned down and discrete ACLs are configured on the Catalyst 2950 to match them, another limitation of the switch would come into play. Specifically, the Catalyst 2950 can only support policing in 1 Mbps increments on FastEthernet ports. Such lax policing would leave a fairly large hole to allow unauthorized traffic that may be mimicking Voice or call signaling to be admitted onto the network.

Due to these limitations, it is not recommended to use a Catalyst 2950 to support an untrusted PC running SoftPhone.

Catalyst 2950—Untrusted Server with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

For the most part, the Catalyst 2950 can support the Untrusted Multi-Application Server + Scavenger Model as illustrated in Figure 2-5. Only the final element of the logical model, namely the policing of all other traffic to 1 Mbps (remarking traffic in excess of this limit to CS1) is not supported on the Catalyst 2950.

The main platform-specific caveats that should be kept in mind when deploying this model on the Catalyst 2950 are the following:

- Non-standard DSCP values are not supported; therefore, Mission-Critical Data traffic cannot be marked to DSCP 25 on Catalyst 2950s (a temporary recommendation during the interim of Cisco's call signaling marking migration from AF31 to CS3); such application traffic can alternatively be marked to the more general class of Transactional Data (AF21), of which they are a subset.
- The **mls qos cos override** interface command must be used to ensure that untrusted CoS values are explicitly set 0 (default).
- The **range** keyword cannot be used in the ACLs being referenced by the class-maps; server-ports should be explicitly defined with a separate access list entry (ACE) per TCP/UDP port.
- User-defined masks must be consistent for all ACLs being referenced by class maps (if filtering is being done against TCP/UDP ports, then all Access Control Entries (ACEs) should be set to filter by TCP/UDP ports, as opposed to some ACEs filtering by ports and others by subnet or host addresses).
- System-defined masks (such as **permit ip any any**) cannot be used in conjunction with user-defined masks (such as **permit tcp any any eq 3200**) within the same policy map; therefore, if some traffic is being matched against TCP/UDP ports, then a final ACL cannot be used to match all other traffic via a **permit ip any any** statement).
- The Catalyst 2950 IOS implementation of MQC's class-default does not (at the time of writing) function compatibly with mainline IOS; class-default should apply a policy to all other traffic not explicitly defined, but testing has shown that this is not the case.

Example 2-3 Catalyst 2950—Untrusted Multi-Application Server with Scavenger-Class QoS Model Configuration

```
CAT2950(config)#class-map SAP
CAT2950(config-cmap)# match access-group name SAP
CAT2950(config-cmap)#class-map LOTUS
CAT2950(config-cmap)# match access-group name LOTUS
CAT2950(config-cmap)#class-map IMAP
CAT2950(config-cmap)# match access-group name IMAP
CAT2950(config-cmap)#exit
CAT2950(config)#
CAT2950(config)#policy-map UNTRUSTED-SERVER
CAT2950(config-pmap)#  class SAP
CAT2950(config-pmap-c)#    set ip dscp 18                    ! DSCP 25 is not supported
CAT2950(config-pmap-c)#    police 15000000 8192 exceed-action dscp 8
    ! Out-of-profile Mission-Critical is marked down to Scavenger (CS1)
CAT2950(config-pmap-c)#  class LOTUS
CAT2950(config-pmap-c)#    set ip dscp 18                    ! Transactional is marked AF21
CAT2950(config-pmap-c)#    police 35000000 8192 exceed-action dscp 8
    ! Out-of-profile Transactional Data is marked down to Scavenger (CS1)
CAT2950(config-pmap-c)#  class IMAP
CAT2950(config-pmap-c)#    set ip dscp 10                    ! Bulk Data is marked AF11
CAT2950(config-pmap-c)#    police 50000000 8192 exceed-action dscp 8
    ! Out-of-profile Bulk Data is marked down to Scavenger (CS1)
CAT2950(config-pmap-c)#exit
CAT2950(config-pmap)#exit
CAT2950(config)#
CAT2950(config)#interface FastEthernet0/1
CAT2950(config-if)# mls qos cos override                ! Untrusted CoS is remarked to 0
CAT2950(config-if)# service-policy input UNTRUSTED-SERVER
CAT2950(config-if)#exit
CAT2950(config)#
CAT2950(config)#ip access list  extended SAP
CAT2950(config-ext-nacl)# permit tcp any eq 3200 any
CAT2950(config-ext-nacl)# permit tcp any eq 3201 any
CAT2950(config-ext-nacl)# permit tcp any eq 3202 any
CAT2950(config-ext-nacl)# permit tcp any eq 3203 any
CAT2950(config-ext-nacl)# permit tcp any eq 3600 any
CAT2950(config-ext-nacl)#
CAT2950(config-ext-nacl)#ip access list  extended LOTUS
CAT2950(config-ext-nacl)# permit tcp any eq 1352 any
CAT2950(config-ext-nacl)#
CAT2950(config-ext-nacl)#ip access list  extended IMAP
CAT2950(config-ext-nacl)# permit tcp any eq 143 any
CAT2950(config-ext-nacl)# permit tcp any eq 220 any
CAT2950(config-ext-nacl)#end
CAT2950#
```

Catalyst MLS QoS Verification Commands

This section includes the following Catalyst MLS verification commands:

- show mls qos interface policers
- show class-map and show policy-map
- show mls masks qos

show mls qos interface policers

The **show mls qos interface policers** verification command reports all configured policers attached to the specified interface.

In the following example, the policers defined for Mission-Critical data, Transactional Data and Bulk data which are applied to FastEthernet 0/1 are confirmed.

Example 2-4 Show MLS QoS Interface Policers Verification of a Switchport Connected to an Untrusted Multi-Application Server

```
CAT2950#show mls qos interface FastEthernet0/1 policers
FastEthernet0/1
policymap=UNTRUSTED-SERVER
type=Single rate=15000000, burst=8192               ! Mission-Critical Data Policer
type=Single rate=35000000, burst=8192               ! Transactional Data Policer
type=Single rate=50000000, burst=8192               ! Bulk Data Policer
CAT2950#
```

show class-map and show policy-map

The **show class-map** and **show policy-map** verification commands will report the class-map and policy-maps that have been globally configured (regardless of whether or not they've been applied to an interface).

In the following example, the class-maps for SAP, LOTUS and IMAP are displayed, as is the policy-map UNTRUSTED-SERVER that is referencing these.

Example 2-5 Show Class-Map and Show Policy-Map Verification of a Switch Connected to an Untrusted Multi-Application Server

```
CAT2950#show class-map
 Class Map match-all IMAP (id 3)
   Match access-group name IMAP

 Class Map match-any class-default (id 0)
   Match any
 Class Map match-all SAP (id 1)
   Match access-group name SAP

 Class Map match-all LOTUS (id 2)
   Match access-group name LOTUS

CAT2950#show policy-map
 Policy Map UNTRUSTED-SERVER
  class  SAP
   set ip dscp 18
   police 15000000 8192 exceed-action dscp 8
  class  LOTUS
   set ip dscp 18
   police 35000000 8192 exceed-action dscp 8
  class  IMAP
   set ip dscp 10
   police 50000000 8192 exceed-action dscp 8
CAT2950#
```

show mls masks qos

The **show mls masks qos** verification command is helpful in keeping track of the number of user-defined or system-defined masks that are being applied by access list entries that are referenced by MQC class-maps.

In the example below, the ACEs being referenced by QoS policies are using IP protocol masks, including (TCP/UDP) source ports.

Example 2-6 Show MLS Masks QoS Verification of an Untrusted Multi-Application Server Model

```
CAT2950#show mls masks qos
Mask1
    Type : qos
    Fields : ip-proto, src-port
    Policymap : UNTRUSTED-SERVER
        Interfaces : Fa0/1
CAT2950#
```

Catalyst 2950—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

When configuring an access-switch to trust/conditionally-trust CoS, then the default mapping for CoS 5 should be adjusted to point to DSCP EF (46), instead of DSCP CS5 (40). This modification is shown below.

Example 2-7 Catalyst 2950—CoS-to-DSCP Marking Modification for Voice

```
CAT2950(config)#mls qos map cos-dscp 0 8 16 24 32 46 48 56          ! Maps CoS 5 to EF
CAT2950(config)#
```

Note Adjusting the default CoS-to-DSCP mapping for call signaling (which formerly was mapped from CoS 3 to DSCP AF31/26) is no longer required. This is because the default mapping of CoS 3 points to DSCP CS3 (24), which is the call signaling marking that all Cisco IP Telephony devices markings will migrate to.

Catalyst MLS QoS Verification Commands

The **show mls qos map [cos-dscp | dscp-cos]** verification command returns the DSCP-to-CoS and CoS-to-DSCP mappings. These mappings may be either the default mappings or manually configured overrides.

In the example below, the default mapping for CoS 5 (DSCP CS5) has been modified to point to DSCP EF instead.

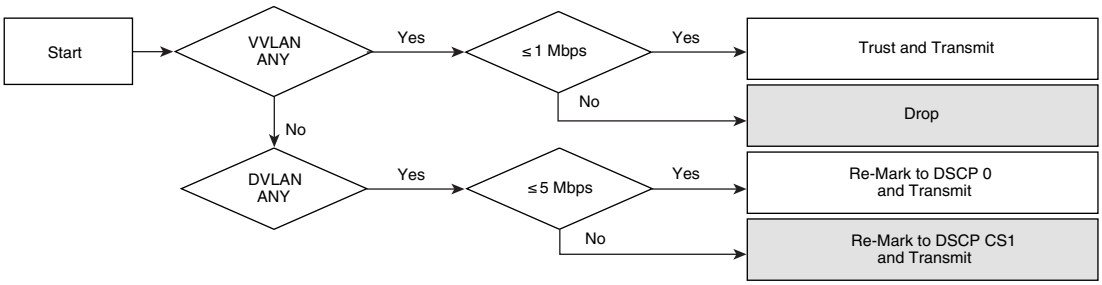
Example 2-8 Show MLS QoS Map Verification for a Catalyst 2950 Switch

```
CAT2950#show mls qos map
   Dscp-cos map:
       dscp:   0  8 10 16 18 24 26 32 34 40 46 48 56
     ------------------------------------------------
```

```
        cos:   0  1  1  2  2  3  3  4  4  5  5  6  7
   Cos-dscp map:
        cos:   0  1  2  3  4  5  6  7
     --------------------------------
       dscp:   0  8 16 24 32 46 48 56                    ! CoS 5 is now mapped to DSCP EF
CAT2950#
```

The Catalyst 2950's hardware policers lack the granularity to implement the Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model, as illustrated in Figure 2-8. However, they can implement a simplified version of this model, as shown in Figure 2-11.

Figure 2-11 Catalyst 2950—Conditionally-Trusted Endpoint Policing: IP Phone + PC with Scavenger-Class QoS (Basic) Model



It should be kept in mind that the coarse granularity of the Catalyst 2950's policers (which are configured in 1 Mbps minimum increments on FastEthernet interfaces) could potentially allow up to 1 Mbps of traffic mimicking legitimate voice traffic per conditionally-trusted switchport.

The configuration for configuring a switchport to conditionally trust an IP Phone that has a PC connected to it, with Scavenger-class QoS, is shown below.

Example 2-9 Catalyst 2950—Conditionally Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model

```
CAT2950(config)#mls qos map cos-dscp 0 8 16 24 32 46 48 56          ! Maps CoS 5 to EF
CAT2950(config)#
CAT2950(config)#class-map VVLAN-ANY
CAT2950(config-cmap)# match access-group name VVLAN-ANY
CAT2950(config-cmap)#class-map DVLAN-ANY
CAT2950(config-cmap)# match access-group name DVLAN-ANY
CAT2950(config-cmap)#exit
CAT2950(config)#
CAT2950(config)#policy-map IPPHONE+PC
CAT2950(config-pmap)#  class VVLAN-ANY
CAT2950(config-pmap-c)#    police 1000000 8192 exceed-action drop
    ! Out-of-profile traffic from the VVLAN is dropped
CAT2950(config-pmap-c)#  class DVLAN-ANY
CAT2950(config-pmap-c)#    set ip dscp 0
    ! Optional remarking in case the trust boundary is compromised
CAT2950(config-pmap-c)#    police 5000000 8192 exceed-action dscp 8
    ! Out-of-profile data traffic is marked down to Scavenger
CAT2950(config-pmap-c)#exit
CAT2950(config-pmap)#exit
CAT2950(config)#
CAT2950(config)#
CAT2950(config)#interface FastEthernet0/1
CAT2950(config-if)# switchport access vlan 10
CAT2950(config-if)# switchport voice vlan 110
CAT2950(config-if)# mls qos trust device cisco-phone               ! Conditional trust
```

```
CAT2950(config-if)# mls qos trust cos                          ! Trust CoS from IP Phone
CAT2950(config-if)# service-policy input IPPHONE+PC               ! Policing policy
CAT2950(config-if)#exit
CAT2950(config)#
CAT2950(config)#ip access list  standard VVLAN-ANY
CAT2950(config-std-nacl)# permit 10.1.110.0 0.0.0.255             ! VVLAN subnet
CAT2950(config-std-nacl)#
CAT2950(config-std-nacl)#ip access list  standard DVLAN-ANY
CAT2950(config-std-nacl)# permit 10.1.10.0 0.0.0.255              ! DVLAN subnet
CAT2950(config-std-nacl)#end
CAT2950#
```

Other Catalyst MLS QoS verification commands include the following:

- show mls qos interface
- show mls qos interface policers
- show mls qos map
- show class-map
- show policy-map
- show mls masks qos

Catalyst 2950—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model

The advanced model of the Conditionally-Trusted IP Phone + PC with Scavenger-class QoS, as shown in Figure 2-9 and Figure 2-10, cannot be supported on the Catalyst 2950 because of the previously discussed caveats and limitations of the Catalyst 2950, including the maximum number of policers supported per FE interface, the overly coarse policer granularity, the inability to mix user-defined masks with system-defined masks and other constraints.

Catalyst 2950—Queuing

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

The Catalyst 2950 can be configured to operate in a 4Q1T mode or in a 1P3Q1T mode (with Queue 4 being configured as a strict-priority queue); the 1P3Q1T mode is recommended for converged networks.

The strict priority queue is enabled by configuring the fourth queue's weight parameter, as defined in the **wrr-queue bandwidth** command, to be 0 (as shown in the example below).

The remaining bandwidth is allocated the other queues according to their defined weights. To allocate remaining bandwidths of 5%, 25% and 70% to Queues 1, 2 and 3, weights of 5, 25, 70 can be assigned these queues, respectively. The logic of these bandwidth allocations recommendations will be discussed in more detail momentarily.

Note Alternatively, as these weights are relative, they can be reduced by dividing each weight by the lowest common denominator (in this case 5) to arrive at queue weights of 1, 5 and 14 for Queues 1, 2 and 3 (respectively). Reduction is strictly optional and makes no difference to the servicing of the queues. Many network administrators tend to prefer defining bandwidth allocation ratios as percentages, and so bandwidth weight ratios aren't reduced in this design chapter.

So far, Campus QoS designs have been presented for the first half of the DoS/worm mitigation strategy discussed at the beginning of this chapter, namely, designs for access layer policers to mark down out-of-profile traffic to the Scavenger class PHB of CS1.

The second vital component of this strategy is to map Scavenger class traffic into a "less-than Best-Effort" queuing structure, ensuring that all other traffic will be serviced ahead of it in the event of congestion.

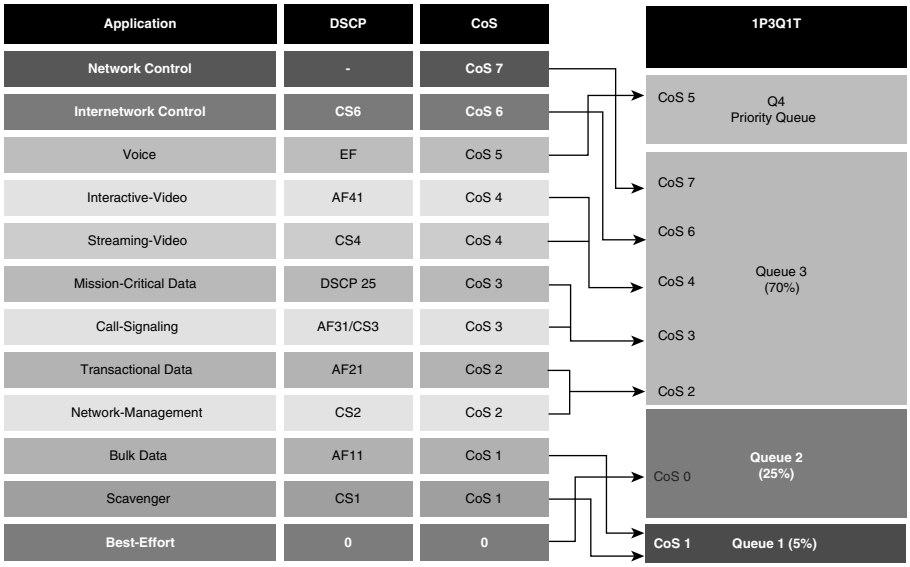
The Catalyst 2950, like most Catalyst platforms, supports the mapping of CoS values into queues. The CoS value that corresponds to Scavenger (DSCP CS1) is CoS 1; this CoS value is shared with Bulk data (DSCP AF11). Therefore, a small amount of bandwidth (5%) is allocated to the "less-than Best-Effort" queue: Q1. Q1 will thus service legitimate Bulk traffic but will constrain out-of-profile Scavenger traffic—which may be the result of a DoS/worm attack—to a small amount (<5%), in the event of congestion.

The next queue, Q2, is then assigned to service Best Effort traffic. A previously discussed design principle regarding Best Effort bandwidth allocation is to allocate approximately 25% of a link's bandwidth to service Best Effort traffic. In this manner the sheer volume of traffic that defaults to Best Effort will continue getting adequate bandwidth, both in the event of momentary campus congestion (due to bursts in the amount of legitimate traffic) and even in the case of a DoS/worm attack.

Preferential applications, such as Transactional Data, Mission-Critical Data, Call-Signaling, Network/Internetwork Control and Management, as well as both Interactive and Streaming Video will be serviced by Q3. Q3 is allocated 70% of the remaining bandwidth (after the PQ has serviced its voice traffic).

The recommended 1P3Q1T queuing model for the Catalyst 2950, along with CoS-to-Queue assignments is illustrated in Figure 2-12.

Figure 2-12 Catalyst 2950 1P3Q1T Queuing Model



The configuration of the priority queue (Q4) and the bandwidth allocations for the remaining queues (Q1, Q2 & Q3) are shown below.

Example 2-10 Catalyst 2950 Scheduling Configuration—1P3Q1T Example

```
CAT2950(config)#wrr-queue bandwidth 5 25 70 0                ! Q1-5%, Q2-25%, Q3-70%, Q4=PQ
CAT2950(config)#
```

Catalyst MLS QoS Verification Commands

This section includes the following commands:

- show wrr-queue bandwidth
- show wrr-queue cos-map

show wrr-queue bandwidth

The **show wrr-queue bandwidth** verification command displays the weights that have been assigned to the queues. If the command returns a value of 0 for the weight of the fourth queue, this indicates that the scheduler is operating in 1P3Q1T mode, with Q4 being the strict-priority queue.

In the following example, the scheduler has been configured for 1P3Q1T queuing: Q1 gets 5% of the remaining bandwidth (after the priority queue has been fully-serviced), Q2 gets 25% and Q3 gets 70%.

Example 2-11 Show WRR-Queue Bandwidth Verification for a Catalyst 2950 Switch

```
CAT2950#show wrr-queue bandwidth
WRR Queue   :   1   2   3   4
Bandwidth   :   5  25  70   0              ! Q1 gets 5%, Q2 gets 25%, Q3 gets 70%, Q4 is PQ
CAT2950#
```

The CoS-to-Queue mapping configuration for the Catalyst 2950 is shown below.

Example 2-12 Catalyst 2950 CoS-to-Queue Mapping Example

```
CAT2950(config)#wrr-queue cos-map 1 1                ! Scavenger/Bulk is assigned to Q1
CAT2950(config)#wrr-queue cos-map 2 0                ! Best Effort is assigned to Q2
CAT2950(config)#wrr-queue cos-map 3 2 3 4 6 7        ! CoS 2,3,4,6,7 are assigned to Q3
CAT2950(config)#wrr-queue cos-map 4 5                ! VoIP is assigned to Q4 (PQ)
CAT2950(config)#
```

show wrr-queue cos-map

The **show wrr-queue cos-map** verification command displayst he queue that each CoS value has been assigned to.

In the example below, CoS 0 (Best Effort) is assigned to Q2 and CoS 1 (Scavenger) has been assigned to Q1. CoS values 2, 3, 4, 6 and 7 have all been assigned to Q3 and CoS 5 (Voice) has been assigned to the priority-queue, Q4.

Example 2-13 Show WRR-Queue CoS Map Verification for a Catalyst 2950 Switch

```
CAT2950#show wrr-queue cos-map
CoS Value      :  0  1  2  3  4  5  6  7
Priority Queue :  2  1  3  3  3  4  3  3

CAT2950#
```

Catalyst 3550—QoS Considerations and Design

This section includes the following topics:

- Catalyst 3550—Trusted Endpoint Model
- Catalyst 3550—AutoQoS VoIP Model
- Catalyst 3550—Untrusted PC + SoftPhone with Scavenger-Class QoS Model
- Catalyst 3550—Untrusted Server with Scavenger-Class QoS Model
- Catalyst 3500—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model
- Catalyst 3550—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model
- Catalyst 3550—Queuing and Dropping

The Catalyst 3550 supports IP routing and so may be found in either the access or distribution layer of the campus.

Regarding QoS, the Catalyst 3550 supports a richer feature set than the Catalyst 2950, including an advanced policing feature that is ideal for out-of-profile policing, namely, per-port/per-VLAN policing. The access layer design options and distribution layer design recommendations for a Catalyst 3550 are shown in Figure 2-13 and Figure 2-14, respectively.

Figure 2-13 Access Layer Catalyst 3550 QoS Design Options

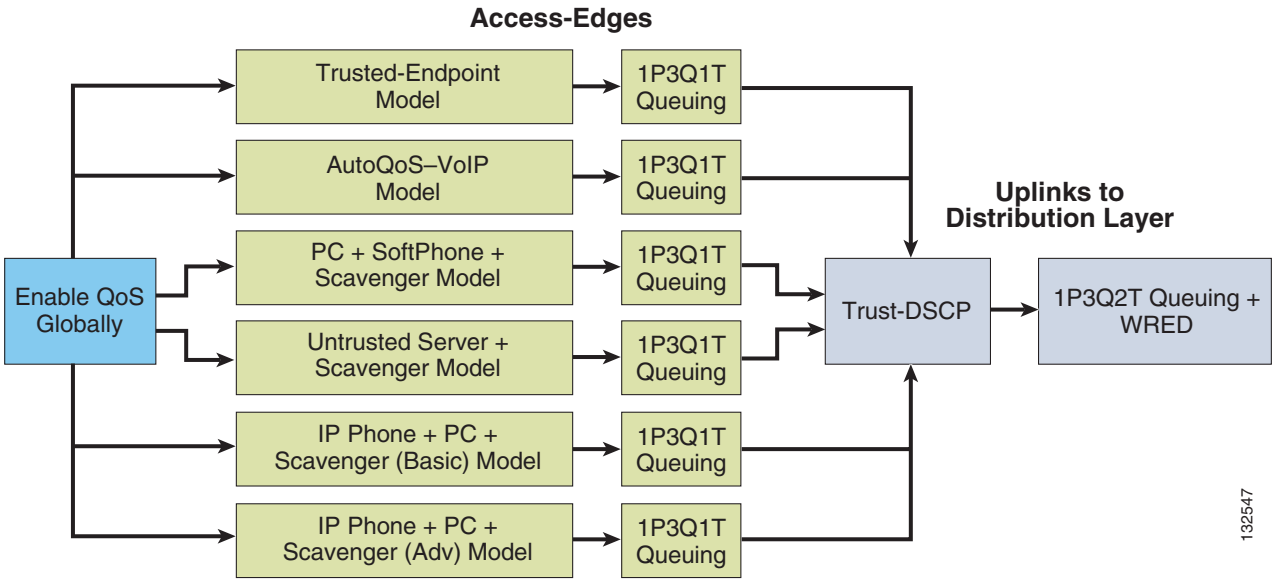
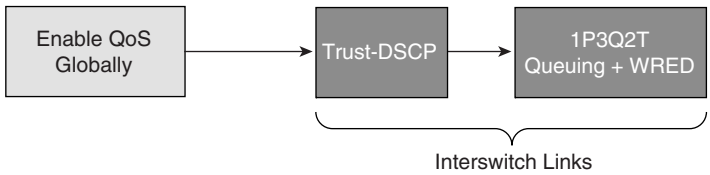


Figure 2-14 Distribution Layer Catalyst 3550 QoS Design



An important point to remember about the Catalyst 3550 is that QoS is disabled by default and must be enabled globally for configured policies to become effective. While QoS is disabled, all frames/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 (by default) set to 0 (which is equivalent to an untrusted state on all ports). The example below shows how to verify if QoS has been enabled or not and also how it can be globally enabled.

Example 2-14 Enabling QoS Globally on the Catalyst 3550

```
CAT3550#show mls qos
QoS is disabled                                          ! By default QoS is disabled

CAT3550#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
CAT3550(config)#mls qos                                  ! Enables QoS globally for the Cat3550
CAT3550(config)#exit
CAT3550#

CAT3550#show mls qos
QoS is enabled                                           ! Verifies that QoS is enabled globally

CAT3550#
```

Note Enabling QoS in the Catalyst 3550 may (depending on the software version) require the disabling of IEEE 802.3X flow control on all interfaces (if enabled). Flowcontrol can be disabled on an interface by using the interface commands: **flowcontrol receive off** and **flowcontrol send off**.

Catalyst 3550—Trusted Endpoint Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

Configuring a Catalyst 3550 switchport to trust an endpoint is identical to the configuration required on a Catalyst 2950, provided that QoS has been globally enabled on the Catalyst 3550. Configuration is shown below.

Example 2-15 Catalyst 3550—Trusted Endpoint

```
CAT3550(config)#interface FastEthernet0/1
CAT3550(config-if)#mls qos trust dscp
```

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for the trusted endpoint model include the following:

- show mls qos
- show mls qos interface

Catalyst 3550—AutoQoS VoIP Model

The Catalyst 3550 supports AutoQoS VoIP with the following keyword options:

- **auto qos voip cisco-phone**
- **auto qos voip cisco-softphone**
- **auto qos voip trust**

When you enable AutoQoS VoIP on the Catalyst 3550 by using the **auto qos voip cisco-phone**, the **auto qos voip cisco-softphone**, or the **auto qos voip trust interface** configuration command, the switch automatically generates a QoS configuration based on the traffic type and ingress packet label and applies the commands listed in Table 2-2 to the interface.

Table 2-2 Catalyst 3550 Auto-QoS Generated Configuration

Description	Automatically Generated Command
The switch automatically enables standard QoS and configures the CoS-to-DSCP map (maps CoS values in incoming packets to a DSCP value).	`C3550(config)# mls qos` `C3550(config)# mls qos map cos-dscp 0 8 16 26 32 46 48 56`

If 10/100 Ethernet ports are present, the switch automatically configures the buffer size of the minimum-reserve levels 5, 6, 7, and 8:
• Level 5 can hold 170 packets.
• Level 6 can hold 85 packets.
• Level 7 can hold 51 packets.
• Level 8 can hold 34 packets.
	`C3550(config)# mls qos min-reserve 5 170`
	`C3550(config)# mls qos min-reserve 6 85`
	`C3550(config)# mls qos min-reserve 7 51`
	`C3550(config)# mls qos min-reserve 8 34`

If you entered 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 or to trust the DSCP value received in the packet on a routed port.	`C3550(config-if)# mls qos trust cos` `C3550(config-if)# mls qos trust dscp`
If you entered the **auto qos voip cisco-phone** command, the switch automatically enables the trusted boundary feature which uses the CDP to detect the presence or absence of a Cisco IP Phone.	`C3550(config-if)# mls qos trust device cisco-phone`

If you entered the **auto qos voip cisco-softphone** command, the switch automatically creates class maps and policy maps.	`C3550(config)# mls qos map policed-dscp 24 26 46 to 0`
	`C3550(config)# class-map match-all AutoQoS-VoIP-RTP-Trust`
	`C3550(config-cmap)# match ip dscp 46`
	`C3550(config)# class-map match-all AutoQoS-VoIP-Control-Trust`
	`C3550(config-cmap)# match ip dscp 24 26`
	`C3550(config)# policy-map AutoQoS-Police-SoftPhone`
	`C3550(config-pmap)# class AutoQoS-VoIP-RTP-Trust`
	`C3550(config-pmap-c)# set dscp 46`
	`C3550(config-pmap-c)# police 320000 8000 exceed-action policed-dscp-transmit`
	`C3550(config-pmap)# class AutoQoS-VoIP-Control-Trust`
	`C3550(config-pmap-c)# set dscp 24`
	`C3550(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit`

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.	`C3550(config-if)# service-policy input AutoQoS-Police-SoftPhone`

Table 2-2 Catalyst 3550 Auto-QoS Generated Configuration

The switch automatically assigns egress queue usage on this interface. The switch enables the egress expedite queue and assigns WRR weights to queues 1, 2, and 3. (The lowest value for a WRR queue is 1and is the recommended setting for Q4 when it is operating as an expedite queue.) The switch configures the CoS-to-egress-queue map: • CoS values 0 and 1 select queue 1. • CoS values 2 and 4 select queue 2. • CoS values 3, 6, and 7 select queue 3. • CoS value 5 selects queue 4 (expedite queue). Because the expedite queue (queue 4) contains the VoIP data traffic, the queue is serviced until empty.	C3550(config-if)# **wrr-queue bandwidth 10 20 70 1** C3550(config-if)# **no wrr-queue cos-map** C3550(config-if)# **wrr-queue cos-map 1 0 1** C3550(config-if)# **wrr-queue cos-map 2 2 4** C3550(config-if)# **wrr-queue cos-map 3 3 6 7** C3550(config-if)# **wrr-queue cos-map 4 5** C3550(config-if)# **priority-queue out**
On Gigabit-capable Ethernet ports only, the switch automatically configures the ratio of the sizes of the WRR egress queues: • Queue 1 is 50 percent. • Queue 2 is 25 percent. • Queue 3 is 15 percent. • Queue 4 is 10 percent.	C3550(config-if)# **wrr-queue queue-limit 50 25 15 10**
On 10/100 Ethernet ports only, the switch automatically configures minimum-reserve levels for the egress queues: • Queue 1 selects the minimum-reserve level 5. • Queue 2 selects the minimum-reserve level 6. • Queue 3 selects the minimum-reserve level 7. • Queue 4 selects the minimum-reserve level 8.	C3550(config-if)# **wrr-queue min-reserve 1 5** C3550(config-if)# **wrr-queue min-reserve 2 6** C3550(config-if)# **wrr-queue min-reserve 3 7** C3550(config-if)# **wrr-queue min-reserve 4 8**

Catalyst 3550—Untrusted PC + SoftPhone with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

Unlike the Catalyst 2950, the Catalyst 3550 has all the necessary QoS features to support and enforce the Untrusted PC + SoftPhone + Scavenger model, as illustrated in Figure 2-4. The Catalyst 3550 configuration for this access edge model is shown below.

Example 2-16 Catalyst 3550—Untrusted PC + SoftPhone + Scavenger Model Configuration

```
CAT3550(config)#mls qos map policed-dscp  0 24 46 to 8
    ! Excess traffic marked 0 or CS3 or EF will be remarked to CS1
CAT3550(config)#
CAT3550(config)#class-map match-all SOFTPHONE-VOICE
CAT3550(config-cmap)#  match access-group name SOFTPHONE-VOICE
CAT3550(config-cmap)#class-map match-all SOFTPHONE-SIGNALING
CAT3550(config-cmap)#  match access-group name SOFTPHONE-SIGNALING
CAT3550(config-cmap)#exit
CAT3550(config)#
CAT3550(config)#policy-map SOFTPHONE-PC
CAT3550(config-pmap)#class SOFTPHONE-VOICE
CAT3550(config-pmap-c)# set ip dscp 46              ! Softphone VoIP is marked to DSCP EF
CAT3550(config-pmap-c)# police 128000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile SoftPhone voice traffic is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class SOFTPHONE-SIGNALING
CAT3550(config-pmap-c)# set ip dscp 24                  ! Signaling is marked to DSCP CS3
CAT3550(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile Signaling traffic is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class class-default
CAT3550(config-pmap-c)# set ip dscp 0
CAT3550(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)# exit
CAT3550(config-pmap)#exit
CAT3550(config)#
CAT3550(config)#interface FastEthernet0/1
CAT3550(config-if)# service-policy input SOFTPHONE-PC           ! Applies policy to int
CAT3550(config-if)#exit
CAT3550(config)#
CAT3550(config)#ip access list  extended SOFTPHONE-VOICE
CAT3550(config-ext-nacl)# permit udp any any range 16384 32767             ! VoIP ports
CAT3550(config-ext-nacl)#
CAT3550(config-ext-nacl)#ip access list  extended SOFTPHONE-SIGNALING
CAT3550(config-ext-nacl)# permit tcp any any range 2000 2002               ! SCCP ports
CAT3550(config-ext-nacl)#end
CAT3550#
```

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for untrusted PC + SoftPhone +Scavenger model include the following:

- show mls qos

- show mls qos map
- show mls qos interface
- show mls qos interface policers
- show mls qos statistics
- show class-map
- show policy-map
- show policy interface

show mls qos interface statistics

The **show mls qos interface statistics** verification command reports dynamic counters for a given policy, including how many packets were classified and policed by the policy.

In the example below, untrusted packets from the PC are classified and policed according to the limits shown in Figure 2-4 for the Untrusted PC + SoftPhone + Scavenger access edge endpoint policing model.

Example 2-17 Show MLS QoS Interface Statistics Verification of a Catalyst 3550 Switchport Connected to an Untrusted PC + SoftPhone with Scavenger-Class QoS

```
CAT3550#show mls qos interface FastEthernet0/1 statistics
FastEthernet0/1
Ingress
  dscp: incoming   no_change  classified policed    dropped (in bytes)
Others: 1275410698 31426318   1243984380 1674978822 0
Egress
  dscp: incoming   no_change  classified policed    dropped (in bytes)
Others: 7271494       n/a       n/a      0          0
CAT3550#
```

show policy interface

The **show policy interface** verification command displays the policy maps (and related classes) that are attached to a given interface.

In this example, a summary of the Untrusted PC + SoftPhone + Scavenger policing policy is shown as applied to FastEthernet0/1.

Example 2-18 Show Policy Interface Verification of a of a Catalyst 3550 Switchport Connected to an Untrusted PC + SoftPhone with Scavenger-Class QoS

```
CAT3550#show policy interface FastEthernet0/1
 FastEthernet0/1
  service-policy input: SOFTPHONE-PC

    class-map: SOFTPHONE-VOICE (match-all)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      match: access-group name SOFTPHONE-VOICEqm_police_inform_feature:
 CLASS_SHOW

    class-map: SOFTPHONE-SIGNALING (match-all)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      match: access-group name SOFTPHONE-SIGNALINGqm_police_inform_feature:
 CLASS_SHOW
```

```
    class-map: class-default (match-any)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      match: any
        0 packets, 0 bytes
        5 minute rate 0 bpsqm_police_inform_feature: CLASS_SHOW

CAT3550#
```

Note Currently, the counters reported via the **show policy interface** command on the Catalyst 3550 are not being incremented (i.e., all counters are currently frozen at zero), as is the case with the mainline IOS version of this command. A bug has been reported which, when fixed, should result in the counters incrementing dynamically. Catalyst 3550 IOS versions tested and affected with this bug include 12.1(19)EA1 a through c and 12.1(20)EA1.

Catalyst 3550—Untrusted Server with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

The Catalyst 3550 fully supports the Untrusted Multi-Application Server with Scavenger-Class QoS model, as depicted in Figure 2-5. The configuration for this model is shown below.

Example 2-19 Catalyst 3550—Untrusted Multi-Application Server with Scavenger-Class QoS Configuration

```
CAT3550(config)#mls qos map policed-dscp  0 10 18 25 to 8
  ! Excess traffic marked 0 or AF11 or AF21 or DSCP 25 will be remarked to CS1
CAT3550(config)#
CAT3550(config)#class-map SAP
CAT3550(config-cmap)# match access-group name SAP
CAT3550(config-cmap)#class-map LOTUS
CAT3550(config-cmap)# match access-group name LOTUS
CAT3550(config-cmap)#class-map IMAP
CAT3550(config-cmap)# match access-group name IMAP
CAT3550(config-cmap)#exit
CAT3550(config)#
CAT3550(config)#policy-map UNTRUSTED-SERVER
CAT3550(config-pmap)#class SAP
CAT3550(config-pmap-c)# set ip dscp 25              ! SAP is marked as Mission-Critical
CAT3550(config-pmap-c)# police 15000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile SAP is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class LOTUS
CAT3550(config-pmap-c)# set ip dscp 18              ! Lotus is marked as Transactional
CAT3550(config-pmap-c)# police 35000000 8000 exceed-action policed-dscp-transmit
! Out-of-profile LOTUS is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class IMAP
CAT3550(config-pmap-c)# set ip dscp 10                  ! IMAP is marked as Bulk Data
CAT3550(config-pmap-c)# police 50000000 8000 exceed-action policed-dscp-transmit
! Out-of-profile IMAP is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class class-default
```

```
CAT3550(config-pmap-c)# set ip dscp 0
CAT3550(config-pmap-c)# police 1000000 8000 exceed-action policed-dscp-transmit
! Out-of-profile excess data traffic is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)# exit
CAT3550(config-pmap)#exit
CAT3550(config)#
CAT3550(config)#interface FastEthernet0/1
CAT3550(config-if)# service-policy input UNTRUSTED-SERVER
CAT3550(config-if)#exit
CAT3550(config)#
CAT3550(config)#ip access list  extended SAP
CAT3550(config-ext-nacl)# permit tcp any range 3200 3203 any
CAT3550(config-ext-nacl)# permit tcp any eq 3600 any
CAT3550(config-ext-nacl)#
CAT3550(config-ext-nacl)#ip access list  extended LOTUS
CAT3550(config-ext-nacl)# permit tcp any eq 1352 any
CAT3550(config-ext-nacl)#
CAT3550(config-ext-nacl)#ip access list  extended IMAP
CAT3550(config-ext-nacl)# permit tcp any eq 143 any
CAT3550(config-ext-nacl)# permit tcp any eq 220 any
CAT3550(config-ext-nacl)#end
CAT3550#
```

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for the Multi-Application Server+ Scavenger model include the following:

- show mls qos
- show mls qos map
- show mls qos interface
- show mls qos interface policers
- show mls qos statistics
- show class-map
- show policy-map
- show policy interface

Catalyst 3500—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

The Catalyst 3550's support of Per-Port/Per-VLAN policing makes for a distinct advantage over other platforms when provisioning a (Basic or Advanced) Conditionally-Trusted Endpoint Model. This is because Per-Port/Per-VLAN policies can be provisioned without having to enter subnet-specific information for each switch. This makes such policies more modular and portable.

In the following example, the VLAN 10 is the DVLAN and VLAN 110 is the VVLAN.

Example 2-20 Catalyst 3550—Conditionally-Trusted IP Phone + PC + Scavenger (Basic) Model Configuration

```
CAT3550(config)#mls qos map cos-dscp 0 8 16 24 32 46 48 56
    ! Modifies CoS-to-DSCP mapping to map CoS 5 to DSCP EF
CAT3550(config)#mls qos map policed-dscp 0 24 to 8
    ! Excess DVLAN & VVLAN traffic will be remarked to Scavenger (CS1)
CAT3550(config)#
CAT3550(config)#
CAT3550(config)#class-map match-all VOICE
CAT3550(config-cmap)#  match ip dscp 46                                ! DSCP EF (voice)
CAT3550(config-cmap)#class-map match-any CALL SIGNALING         ! Need 'match-any' here
CAT3550(config-cmap)#  match ip dscp 26                 ! DSCP AF31 (old Call-Signaling)
CAT3550(config-cmap)#  match ip dscp 24                 ! DSCP CS3 (new Call-Signaling)
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all VVLAN-VOICE
CAT3550(config-cmap)#  match vlan  110                             ! VLAN 110 is VVLAN
CAT3550(config-cmap)#  match class-map VOICE                    ! Matches VVLAN DSCP EF
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all VVLAN-CALL-SIGNALING
CAT3550(config-cmap)#  match vlan  110                             ! VLAN 110 is VVLAN
CAT3550(config-cmap)#  match class-map CALL SIGNALING           !Matches VVLAN AF31/CS3
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all ANY
CAT3550(config-cmap)#  match access-group name ANY                     ! Workaround ACL
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all VVLAN-ANY
CAT3550(config-cmap)#  match vlan  110                             ! VLAN 110 is VVLAN
CAT3550(config-cmap)#  match class-map ANY              ! Matches any other VVLAN traffic
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all DVLAN-ANY
CAT3550(config-cmap)#  match vlan  10                              ! VLAN 10 is DVLAN
CAT3550(config-cmap)#  match class-map ANY                  ! Matches all DVLAN traffic
CAT3550(config-cmap)#
CAT3550(config-cmap)#policy-map IPPHONE+PC-BASIC
CAT3550(config-pmap)#class VVLAN-VOICE
CAT3550(config-pmap-c)# set ip dscp 46                                  ! DSCP EF (Voice)
CAT3550(config-pmap-c)# police 128000 8000 exceed-action drop
    ! Only one voice call is permitted per switchport VVLAN
CAT3550(config-pmap-c)#class VVLAN-CALL-SIGNALING
CAT3550(config-pmap-c)# set ip dscp 24                       ! DSCP CS3 (Call-Signaling)
CAT3550(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile call signaling is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class VVLAN-ANY
CAT3550(config-pmap-c)# set ip dscp 0
CAT3550(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Unauthorized VVLAN traffic is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class DVLAN-ANY
CAT3550(config-pmap-c)# set ip dscp 0
CAT3550(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)# exit
CAT3550(config-pmap)#exit
CAT3550(config)#
CAT3550(config)#interface FastEthernet0/1
CAT3550(config-if)# switchport access vlan 10                                ! DVLAN
CAT3550(config-if)# switchport voice vlan 110                                ! VVLAN
CAT3550(config-if)# mls qos trust device cisco-phone                ! Conditional Trust
CAT3550(config-if)# service-policy input IPPHONE+PC-BASIC              ! Attaches policy
CAT3550(config-if)#exit
CAT3550(config)#
```

```
CAT3550(config)#
CAT3550(config)#ip access list  standard ANY                          ! Workaround ACL
CAT3550(config-std-nacl)# permit any
CAT3550(config-std-nacl)#end
CAT3550#
```

Catalyst MLS QoS Verification Commands

The Catalyst MLS QoS verification commands for the Conditionally-Trusted IP Phone + PC + Scavenger (Basic) Model Configuration include the following:

- show mls qos
- show mls qos map
- show mls qos interface
- show mls qos interface policers
- show mls qos statistics
- show class-map
- show policy-map
- show policy interface

Note While Catalyst 3550 IOS syntax supports the **match any** criteria within a class-map (which the parser allows to be configured in conjunction with a per-VLAN policy), testing with some versions of Catalyst 3500 IOS have revealed a limitation with this function, since it does not match any other traffic on a per-VLAN basis. Hence an explicit access list named ANY has been used in the example above as a workaround.

Catalyst 3550—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

The Conditionally-Trusted IP Phone + PC + Scavenger Advanced model builds on the Basic model by including policers for PC-based video-conferencing (and/or PC SoftPhone), Mission-Critical Data applications, Transactional Data applications, and Bulk Data applications. This model is graphically depicted in Figure 2-9. The Catalyst 3550 can support 8 policers per 10/100 Ethernet port and so can support this Advanced Model.

The Catalyst 3550 configuration for the Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model is shown below. In this example, the same Server-to-Client applications are used as in the Untrusted Multi-Application Server example. However, notice that the Source/Destination ports are reversed for the Client-to-Server direction of traffic flow. Also, due to the limit of the number of

policers per FastEthernet port (8), there is no explicit policer for SoftPhone call signaling traffic; to work around this limitation, SoftPhone call signaling traffic is included in the Mission-Critical Data applications access list .

Example 2-21 Catalyst 3550—Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model Configuration

```
CAT3550(config)#mls qos map cos-dscp 0 8 16 24 32 46 48 56
    ! Modifies CoS-to-DSCP mapping to map CoS 5 to DSCP EF
CAT3550(config)#mls qos map policed-dscp 0 10 18 24 25 34 to 8
! Excess DVLAN traffic marked 0, AF11, AF21, CS3, DSCP 25,
! and AF41 will be remarked to Scavenger (CS1)
CAT3550(config)#
CAT3550(config)#class-map match-all VOICE
CAT3550(config-cmap)#  match ip dscp 46                                  ! DSCP EF (voice)
CAT3550(config-cmap)#class-map match-any CALL SIGNALING          ! Need 'match-any' here
CAT3550(config-cmap)#  match ip dscp 26                 ! DSCP AF31 (old Call-Signaling)
CAT3550(config-cmap)#  match ip dscp 24                 ! DSCP CS3 (new Call-Signaling)
CAT3550(config-cmap)#class-map match-all PC-VIDEO
CAT3550(config-cmap)#  match access-group name PC-VIDEO
CAT3550(config-cmap)#class-map match-all MISSION-CRITICAL-DATA
CAT3550(config-cmap)#  match access-group name MISSION-CRITICAL-DATA
CAT3550(config-cmap)#class-map match-all TRANSACTIONAL-DATA
CAT3550(config-cmap)#  match access-group name TRANSACTIONAL-DATA
CAT3550(config-cmap)#class-map match-all BULK-DATA
CAT3550(config-cmap)#  match access-group name BULK-DATA
CAT3550(config-cmap)#class-map match-all ANY
CAT3550(config-cmap)#  match access-group name ANY                       ! Workaround ACL
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all VVLAN-VOICE
CAT3550(config-cmap)#  match vlan  110                              ! VLAN 110 is VVLAN
CAT3550(config-cmap)#  match class-map VOICE                     ! Matches VVLAN DSCP EF
CAT3550(config-cmap)#class-map match-all VVLAN-CALL-SIGNALING
CAT3550(config-cmap)#  match vlan  110                              ! VLAN 110 is VVLAN
CAT3550(config-cmap)#  match class-map CALL SIGNALING           ! Matches VVLAN AF31/CS3
CAT3550(config-cmap)#class-map match-all VVLAN-ANY
CAT3550(config-cmap)#  match vlan  110                              ! VLAN 110 is VVLAN
CAT3550(config-cmap)#  match class-map ANY              ! Matches any other VVLAN traffic
CAT3550(config-cmap)#
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all DVLAN-PC-VIDEO
CAT3550(config-cmap)#  match vlan  10                                    ! VLAN 10 is DVLAN
CAT3550(config-cmap)#  match class-map PC-VIDEO         ! Matches PC IP/VC or SoftPhone
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all DVLAN-MISSION-CRITICAL-DATA
CAT3550(config-cmap)#  match vlan  10                                    ! VLAN 10 is DVLAN
CAT3550(config-cmap)#  match class-map MISSION-CRITICAL-DATA
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all DVLAN-TRANSACTIONAL-DATA
CAT3550(config-cmap)#  match vlan  10                                    ! VLAN 10 is DVLAN
CAT3550(config-cmap)#  match class-map TRANSACTIONAL-DATA
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all DVLAN-BULK-DATA
CAT3550(config-cmap)#  match vlan  10                                    ! VLAN 10 is DVLAN
CAT3550(config-cmap)#  match class-map BULK-DATA
CAT3550(config-cmap)#
CAT3550(config-cmap)#class-map match-all DVLAN-ANY
CAT3550(config-cmap)#  match vlan  10                                    ! VLAN 10 is DVLAN
CAT3550(config-cmap)#  match class-map ANY              ! Matches all other DVLAN traffic
CAT3550(config-cmap)#
CAT3550(config-cmap)#
CAT3550(config-cmap)#policy-map IPPHONE+PC-ADVANCED
```

```
CAT3550(config-pmap)#class VVLAN-VOICE
CAT3550(config-pmap-c)# set ip dscp 46                                    ! DSCP EF (Voice)
CAT3550(config-pmap-c)# police 128000 8000 exceed-action drop
    ! Only one voice call is permitted per switchport VVLAN
CAT3550(config-pmap-c)#class VVLAN-CALL-SIGNALING
CAT3550(config-pmap-c)# set ip dscp 24                        ! DSCP CS3 (Call-Signaling)
CAT3550(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile call signaling is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class VVLAN-ANY
CAT3550(config-pmap-c)# set ip dscp 0
CAT3550(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
! Unauthorized VVLAN traffic is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class DVLAN-PC-VIDEO
CAT3550(config-pmap-c)# set ip dscp 34                     ! DSCP AF41 (Interactive-Video)
CAT3550(config-pmap-c)# police 500000 8000 exceed-action policed-dscp-transmit
    ! Only one IP/VC stream will be permitted per switchport
CAT3550(config-pmap-c)#class DVLAN-MISSION-CRITICAL-DATA
CAT3550(config-pmap-c)# set ip dscp 25                     ! Interim Mission-Critical Data
CAT3550(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile Mission-Critical Data is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class DVLAN-TRANSACTIONAL-DATA
CAT3550(config-pmap-c)# set ip dscp 18                     ! DSCP AF21 (Transactional Data)
CAT3550(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile Transactional Data is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class DVLAN-BULK-DATA
CAT3550(config-pmap-c)# set ip dscp 10                           ! DSCP AF11 (Bulk Data)
CAT3550(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile Bulk Data is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)#class DVLAN-ANY
CAT3550(config-pmap-c)# set ip dscp 0
CAT3550(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT3550(config-pmap-c)# exit
CAT3550(config-pmap)#exit
CAT3550(config)#
CAT3550(config)#interface FastEthernet0/1
CAT3550(config-if)# switchport access vlan 10                                ! DVLAN
CAT3550(config-if)# switchport voice vlan 110                                ! VVLAN
CAT3550(config-if)# mls qos trust device cisco-phone                ! Conditional Trust
CAT3550(config-if)# service-policy input IPPHONE+PC-ADVANCED           ! Attaches policy
CAT3550(config-if)#exit
CAT3550(config)#
CAT3550(config)#ip access list  standard ANY                              ! Workaround ACL
CAT3550(config-std-nacl)# permit any
CAT3550(config-std-nacl)#
CAT3550(config-std-nacl)#ip access list  extended PC-VIDEO          ! IP/VC or SoftPhone
CAT3550(config-ext-nacl)# permit udp any any range 16384 32767
CAT3550(config-ext-nacl)#
CAT3550(config-ext-nacl)#ip access list  extended MISSION-CRITICAL-DATA
CAT3550(config-ext-nacl)# permit tcp any any range 3200 3203                  ! SAP
CAT3550(config-ext-nacl)# permit tcp any any eq 3600                          ! SAP
CAT3550(config-ext-nacl)# permit tcp any any range 2000 2002             ! SoftPhone SCCP
CAT3550(config-ext-nacl)#
CAT3550(config-ext-nacl)#ip access list  extended TRANSACTIONAL-DATA
CAT3550(config-ext-nacl)# permit tcp any any eq 1352                          ! Lotus
CAT3550(config-ext-nacl)#
CAT3550(config-ext-nacl)#ip access list  extended BULK-DATA
CAT3550(config-ext-nacl)# permit tcp any any eq 143                           ! IMAP
CAT3550(config-ext-nacl)# permit tcp any any eq 220                           ! IMAP
CAT3550(config-ext-nacl)#end
CAT3550#
```

Catalyst MLS QoS Verification Commands

The Catalyst MLS QoS verification commands for Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model include the following:

- show mls qos
- show mls qos map
- show mls qos interface
- show mls qos interface policers
- show mls qos statistics
- show class-map
- show policy-map
- show policy interface

Catalyst 3550—Queuing and Dropping

This section includes the following topics:

- Configuration
- Advanced Tuning Options
- Catalyst MLS QoS Verification Commands

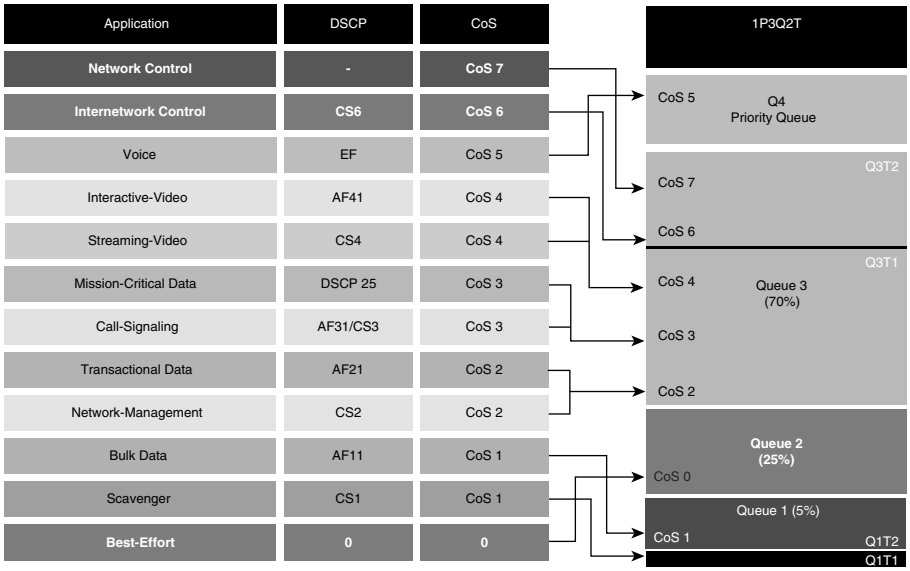
Configuration

Like the Catalyst 2950, the Catalyst 3550 supports a 1P3Q1T queuing model for all ports. GigabitEthernet ports have the additional option of being configured as 1P3Q2T, with either tail-drop or WRED thresholds. However, unlike the Catalyst 2950, the Catalyst 3550 queuing parameters are set on a per-interface basis and not globally. Nonetheless, uniform queuing policies can be expeditiously deployed via the **interface range** configuration command.

The strict-priority queue is enabled on a per-interface basis on the Catalyst 3550 with the **priority-queue out** interface command. Bandwidth is allocated among the remaining queues via the **wrr-queue bandwidth** command. A twist with the Catalyst 3550 is that Queue 4's WRR weight is set to 1 (indicating that it does not participate in the WRR scheduler since it is configured as a strict-priority queue), instead of 0 (as is the case on the Catalyst 2950). Recommended remaining bandwidth allocations (after the PQ has been fully serviced) are 5% for the Scavenger queue (Q1), 25% for the Best-Effort queue (Q2), and 70% for the preferential application queue (Q3).

Following this, CoS 1 (Scavenger/Bulk) would be assigned to Q1; CoS 0 (Best Effort) would be assigned to Q2; CoS values 2 (Transactional Data and Network Management), 3 (call signaling and Mission-Critical Data), 4 (Interactive and Streaming-Video, 6 (Internetwork Control) and 7 (Network Control/Spanning-Tree) would be assigned to Q3; CoS 5 (voice) would be assigned to the strict-priority Q4. These assignments and allocations are illustrated in Figure 2-15 (the thresholds shown in Q1 and Q3 are discussed shortly).

Figure 2-15 Catalyst 3550 1P3Q2T Queuing Model



The interface-mode configuration commands to configure this 1P3Q1T queuing model, for either FastEthernet or GigabitEthernet Catalyst 3550 interfaces, are shown below.

Example 2-22 Catalyst 3550 FastEthernet and/or GigabitEthernet Interface Queuing Configuration—1P3Q1T

```
CAT3550(config)#interface range FastEthernet0/1 - 48
CAT3550(config-if)# wrr-queue bandwidth 5 25 70 1          ! Q1-5% Q2-25% Q3-70% Q4=PQ
CAT3550(config-if)# wrr-queue cos-map 1 1                  ! Assigns Scavenger to Q1
CAT3550(config-if)# wrr-queue cos-map 2 0                  ! Assigns Best Effort to Q2
CAT3550(config-if)# wrr-queue cos-map 3 2 3 4 6 7          ! Assigns CoS 2,3,4,6,7 to Q3
CAT3550(config-if)# wrr-queue cos-map 4 5                  ! Assigns VoIP to Q4 (PQ)
CAT3550(config-if)# priority-queue out                     ! Enables Q4 as PQ
CAT3550(config-if)#exit
CAT3550(config)#
```

Advanced Tuning Options

The Catalyst 3550 offers some advanced "nerd-knob" queuing options on 10/100 interfaces, such as tuning Minimum Reserve Thresholds. However, testing has shown that such tuning, which rarely factors into play, makes a highly-negligible difference at best. Therefore, tuning Minimum Reserve Thresholds is recommended only for advanced network administrators or via automated tools, such as AutoQoS.

Some advanced "nerd-knobs" also exist for GigabitEthernet interfaces. A couple of these advanced tuning options include Queue-Limit tuning and the enabling of WRED thresholds for 1P3Q2T operation. In the event of DoS/worm attacks, the GigabitEthernet uplinks will likely be the first to congest, therefore its worthwhile examining these advanced options.

In the example below, the queue-limits for both GigabitEthernet interfaces are tuned to correspond to the WRR weights of the queues (the bandwidth allocations). This is achieved with the **wrr-queue queue-limit** interface command. However, unlike the WRR weight bandwidth ratio for Q4 (which is set to 1 to indicate that Q4 is a PQ), the queue limit for Q4 needs to be explicitly set to a more representative value, such as 30%.

Note The default queue limits are such that each queue is assigned 25% of the available buffer space. It should be noted that when the queue limits are modified, the queue is temporarily shutdown during the hardware reconfiguration and the switch may drop newly-arrived packets destined to the queue. Thus, it may be advisable not to tune the queue limits on Catalyst 3550 switches already in production networks.

Additionally, WRED is enabled on each (non-priority) queue. This allows for the preferential treatment of Bulk Data (DSCP AF11) over Scavenger (CS1) within Q1, as well as the preferential treatment of Internetworking/Networking protocols (DSCP CS6 and CS7, respectively) over all other applications assigned to Q3. Even though Q2 has only Best Effort traffic assigned to it, enabling WRED on this queue increases the efficiency of TCP applications within this queue during periods of congestion.

A low WRED threshold, such as 40%, can be set for Q1 to aggressively drop Scavenger traffic in order to preferentially service Bulk Data. The WRED thresholds for Q2 and Q3 can be set to higher levels, such as 80%.

By default all DSCP values are mapped to the first WRED threshold of the queue to which their CoS values are assigned. Therefore, only DSCP values that are to be mapped to the second WRED thresholds (of their respective queues) need to be manually configured. In this case, Bulk Data (DSCP AF11/10), Internetwork Control (DSCP CS6/48), and Network Control (DSCP CS7/56) all need to be explicitly mapped to the second WRED threshold via the **wrr-queue dscp-map** interface configuration command.

Note Network control traffic in the campus primarily refers to Spanning Tree Protocol (STP) traffic, such as Bridge Protocol Data Units (BPDUs). While these Layer 2 Ethernet frames are marked CoS 7, they (obviously) do not have any capability to carry Layer 3 DSCP markings. Thus, it may seem moot to map DSCP CS7 (56) to a higher WRED threshold. However, it should be kept in mind that Catalyst switches generate *Internal DSCP* values for all frames (regardless of whether they are carrying IP or not). These Internal DSCP values are used for QoS decisions, such as WRED in this case. Therefore, since STP BPDU frames (marked CoS 7) generate an Internal DSCP value of 56, mapping DSCP 56 to the second threshold of Q3 provides preferential treatment for these important Layer 2 frames.

The configuration for these tuning options, which are available only on GigabitEthernet interfaces on the Catalyst 3550, is shown below.

Example 2-23 Catalyst 3550 GigabitEthernet Interface Queuing and Dropping Configuration—1P3Q2T

```
CAT3550(config)#interface range GigabitEthernet 0/1 - 2
CAT3550(config-if-range)# wrr-queue bandwidth 5 25 70 1
! Q1 gets 5% BW, Q2 gets 25% BW, Q3 gets 70% BW, Q4 is the PQ
CAT3550(config-if-range)# wrr-queue queue-limit 5 25 40 30
    ! Tunes buffers to 5% for Q1, 25% for Q2, 40% for Q3 and 30% for Q4
CAT3550(config-if-range)# wrr-queue random-detect max-threshold 1 40 100
    ! Sets Q1 WRED threshold 1 to 40% and threshold 2 to 100%
CAT3550(config-if-range)# wrr-queue random-detect max-threshold 2 80 100
! Sets Q2 WRED threshold 1 to 80% and threshold 2 to 100%
CAT3550(config-if-range)# wrr-queue random-detect max-threshold 3 80 100
! Sets Q3 WRED threshold 1 to 80% and threshold 2 to 100%
CAT3550(config-if)# wrr-queue cos-map 1 1                  ! Assigns Scavenger to Q1
CAT3550(config-if)# wrr-queue cos-map 2 0                  ! Assigns Best Effort to Q2
CAT3550(config-if)# wrr-queue cos-map 3 2 3 4 6 7          ! Assigns CoS 2,3,4,6,7 to Q3
CAT3550(config-if)# wrr-queue cos-map 4 5                  ! Assigns VoIP to Q4 (PQ)
CAT3550(config-if-range)# wrr-queue dscp-map 2 10 48 56
    ! Maps Bulk Data (10), Routing (48) and Spanning Tree (Internal DSCP 56)
    ! to WRED threshold 2 of their respective queues - all other DSCP values
    ! are mapped (by default) to WRED threshold 1 of their respective queues
CAT3550(config-if)# priority-queue out                     ! Enables Q4 as PQ
```

```
CAT3550(config-if-range)#end
CAT3550#
```

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for queuing and dropping include the following:

- show mls qos interface buffers
- show mls qos interface queueing

show mls qos interface buffers

The **show mls qos interface buffers** verification command displays the queue sizes (as per-queue buffer allocation percentages of the total buffer space). Also the command displays if WRED has been enabled on a queue, and if so, it displays the first and second thresholds (as percentages of the queue's depth).

In the example below, the queue-limits are set to 5%, 25%, 40%, and 30% of the total queuing buffer space for Queues 1 through 4 (respectively). Additionally, WRED is enabled on queues 1 through 3 (but not Q4, as it is the priority queue). The first WRED threshold is set to 5% on Q1 and is set to 80% on queues 2 and 3.

Example 2-24 Show MLS QoS Interface Buffers Verification for a Catalyst 3550 Switch

```
CAT3550#show mls qos interface GigabitEthernet0/1 buffers
GigabitEthernet0/1
Notify Q depth:
qid-size
 1 - 5                          ! Q1 queue-limit is set to 5% of total buffer space
 2 - 25                         ! Q2 queue-limit is set to 25% of total buffer space
 3 - 40                         ! Q3 queue-limit is set to 40% of total buffer space
 4 - 30                         ! Q4 queue-limit is set to 30% of total buffer space
qid WRED thresh1 thresh2
1   ena  40      100           ! WRED is enabled on Q1 - first threshold is set to 40%
2   ena  80      100           ! WRED is enabled on Q2 - first threshold is set to 80%
3   ena  80      100           ! WRED is enabled on Q3 - first threshold is set to 80%
4   dis  100     100           ! WRED is disabled on Q4 (as it is the PQ)

CAT3550#
```

show mls qos interface queueing

The **show mls qos interface queuing** verification command displays the CoS-to-Queuing mappings that have been configured in addition to the bandwidth allocations per queue. On GigabitEthernet interfaces with WRED enabled, the output also includes the DSCP-to-WRED Threshold mappings. This information is displayed in a table form, with the first digit of the decimal DSCP value along the Y-Axis (in rows) and the second digit of the decimal DSCP value along the X-Axis (in columns).

In the example below, the output verifies that the egress expedite queue (priority queue: Q4) is enabled. Also, the WRR Bandwidth Weights show that, of the remaining bandwidth, Q1 is allocated 5%, Q2 is allocated 25%, and Q3 is allocated 70%.

Additionally, the DSCP-to-WRED table verifies that Bulk Data (AF11/10), Internetwork Control (DSCP CS6/48), and Network Control (DSCP CS7/56) are each mapped to the second WRED thresholds (T2) of their respective queues (as determined by the CoS-to-Queue mappings).

Finally, the CoS-to-Queue map shows that CoS 0 (Best Effort) is assigned to Q2, CoS 1 (Scavenger) has been assigned to Q1, CoS values 2, 3, 4, 6 and 7 have all been assigned to Q3, and CoS 5 (Voice) has been assigned to the priority-queue, Q4.

Example 2-25 Show MLS QoS Interface Verification for a Catalyst 3550 Switch

```
CAT3550#show mls qos interface GigabitEthernet0/1 queueing
GigabitEthernet0/1
Egress expedite queue: ena                                    ! Q4 is enabled as a PQ
wrr bandwidth weights:
qid-weights
 1 - 5                                                        ! Q1 is allocated 5%
 2 - 25                                                       ! Q2 is allocated 25%
 3 - 70                                                       ! Q3 is allocated 70&
 4 - 1    when expedite queue is disabled
Dscp-threshold map:
     d1 :  d2 0  1  2  3  4  5  6  7  8  9
     ----------------------------------------
      0 :    01 01 01 01 01 01 01 01 01 01
      1 :    02 01 01 01 01 01 01 01 01 01              ! DSCP 10 is mapped to WRED T2
      2 :    01 01 01 01 01 01 01 01 01 01
      3 :    01 01 01 01 01 01 01 01 01 01
      4 :    01 01 01 01 01 01 01 01 02 01              ! DSCP 48 is mapped to WRED T2
      5 :    01 01 01 01 01 01 02 01 01 01              ! DSCP 56 is mapped to WRED T2
      6 :    01 01 01 01
Cos-queue map:
cos-qid
 0 - 2                    ! Best-Effort is assigned to Q2
 1 - 1                    ! Scavenger and Bulk are assigned to Q1
 2 - 3                    ! Transactional Data and Network Management are assigned to Q3
 3 - 3                    ! Mission-Critical Data and call signaling are assigned to Q3
 4 - 3                    ! Interactive- and Streaming-Video are assigned to Q3
 5 - 4                    ! Voice is assigned to the priority queue: Q4
 6 - 3                    ! Internetwork Control (Routing) is assigned to Q3
 7 - 3                    ! Network Control (Spanning Tree) is assigned to Q3

CAT3550#
```

Catalyst 2970/3560/3750—QoS Considerations and Design

This section includes the following topics:

- Catalyst 2970/3560/3750—Trusted Endpoint Model
- Catalyst 2970/3560/3750—Auto QoS VoIP Model
- Catalyst 2970/3560/3750—Untrusted PC + SoftPhone with Scavenger-Class QoS Model
- Catalyst 2970/3560/3750—Untrusted Server with Scavenger-Class QoS Model
- Catalyst 2970/3560/3750—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model
- Catalyst 2970/3560/3750—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model
- Catalyst 2970/3560/3750—Queuing and Dropping

The Catalyst 2970 does not support Layer 3 routing and, as such, is restricted to the role of an access-layer switch. The 3560 does support Layer 3 routing as well as inline power— a feature that is rarely, if ever, required at the distribution layer—and so is only considered in an access-layer context. The Catalyst 3750 also supports Layer 3 routing and may be found in either the access layer or the distribution layer.

Figure 2-16 shows the QoS design options for access-layer Catalyst 2970s, 3560s, or 3750s, and Figure 2-17 shows the QoS design recommendations for a distribution-layer Catalyst 3750.

Figure 2-16 Access Layer Catalyst 2970/3560/3750 QoS Design Options

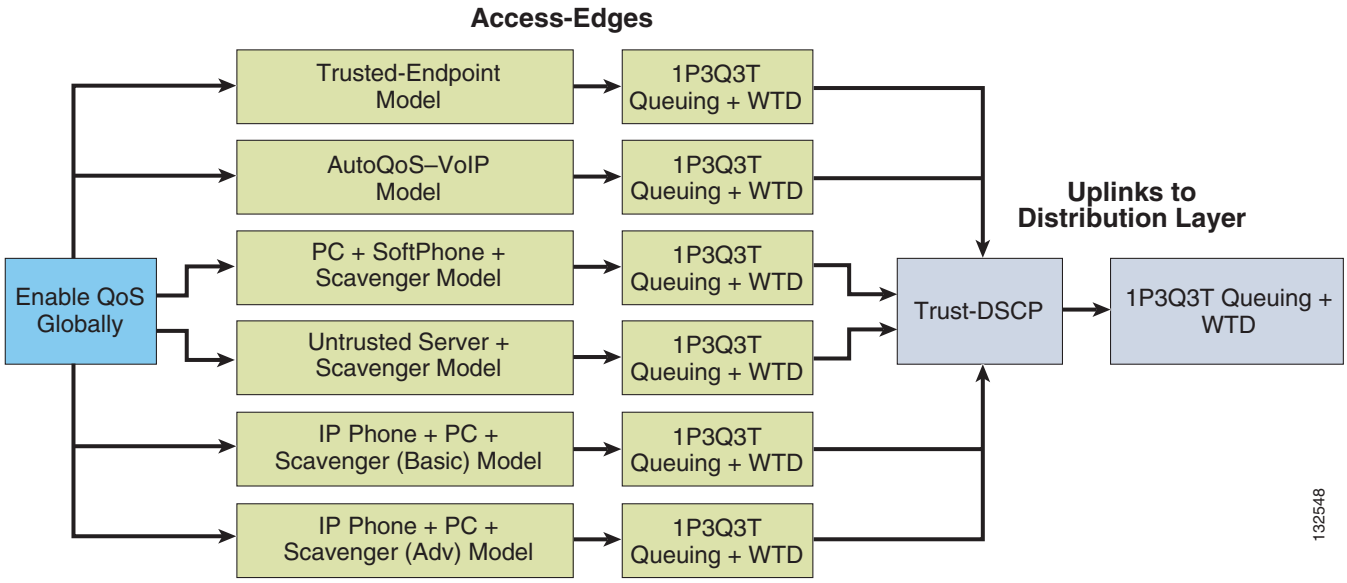
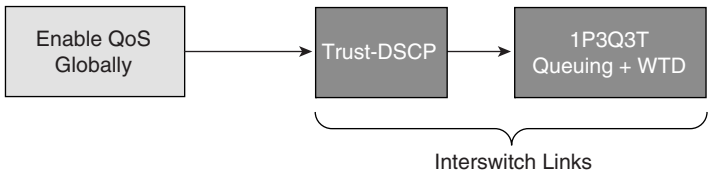


Figure 2-17 Distribution Layer Catalyst 3750 QoS Design



Since the QoS features and configuration syntax are identical for the Catalyst 2970, 3560 and 3750, from a QoS design recommendation perspective, they are subsequently addressed as a single switch.

As with the Catalyst 3550, QoS is globally disabled by default on the Catalyst 2970/3560/3750. While QoS is disabled, all frames/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 (by default) set to 0 (which is equivalent to an untrusted state on all ports).

QoS must be enabled globally for configured policies to become effective. The example below shows how to verify if QoS has been enabled or not and also how it can be globally enabled.

Example 2-26 Enabling QoS Globally on the Catalyst 2970/3560/3750

```
CAT2970#show mls qos
QoS is disabled

CAT2970#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
CAT2970(config)#mls qos
CAT2970(config)#end
CAT2970#

CAT2970#show mls qos
QoS is enabled

CAT2970#
```

Catalyst 2970/3560/3750—Trusted Endpoint Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

The Trusted Endpoint model configuration for the Catalyst 2970/3550 is identical to the switches previously discussed (namely, the Catalyst 2950 and 3550) and is shown below.

Example 2-27 Catalyst 2970/3560/3750—Trusted Endpoint Model Configuration

```
CAT2970(config)#interface GigabitEthernet0/1
CAT2970(config-if)#mls qos trust dscp
```

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for the Trusted Endpoint model include the following:

- show mls qos
- show mls qos interface

Catalyst 2970/3560/3750—Auto QoS VoIP Model

The Catalyst 2970/3560/3750 supports AutoQoS VoIP with the following keyword options:

- **auto qos voip cisco-phone**
- **auto qos voip cisco-softphone**
- **auto qos voip trust**

When you enable AutoQoS VoIP on the Catalyst 2970/3560/3750 by using the **auto qos voip cisco-phone**, the **auto qos voip cisco-softphone**, or the **auto qos voip trust** interface configuration command, the switch automatically generates a QoS configuration based on the traffic type and ingress packet label and applies the commands listed in Table 2-3 to the interface.

Table 2-3 Catalyst 2970/3560/3750 Auto-QoS Generated Configuration

Description	Automatically Generated Command
The switch automatically enables standard QoS and configures the CoS-to-DSCP map (maps CoS values in incoming packets to a DSCP value).	C2970(config)# **mls qos** C2970(config)# **mls qos map cos-dscp 0 8 16 26 32 46 48 56**

Table 2-3 Catalyst 2970/3560/3750 Auto-QoS Generated Configuration

The switch automatically maps CoS values to an ingress queue and to a threshold ID.	C2970(config)# **no mls qos srr-queue input cos-map** C2970(config)# **mls qos srr-queue input cos-map queue 1 threshold 3 0** C2970(config)# **mls qos srr-queue input cos-map queue 1 threshold 2 1** C2970(config)# **mls qos srr-queue input cos-map queue 2 threshold 1 2** C2970(config)# **mls qos srr-queue input cos-map queue 2 threshold 2 4 6 7** C2970(config)# **mls qos srr-queue input cos-map queue 2 threshold 3 3 5**
The switch automatically maps CoS values to an egress queue and to a threshold ID.	C2970(config)# **no mls qos srr-queue output cos-map** C2970(config)# **mls qos srr-queue output cos-map queue 1 threshold 3 5** C2970(config)# **mls qos srr-queue output cos-map queue 2 threshold 3 3 6 7** C2970(config)# **mls qos srr-queue output cos-map queue 3 threshold 3 2 4** C2970(config)# **mls qos srr-queue output cos-map queue 4 threshold 2 1** C2970(config)# **mls qos srr-queue output cos-map queue 4 threshold 3 0**
The switch automatically maps DSCP values to an ingress queue and to a threshold ID.	C2970(config)# **no mls qos srr-queue input dscp-map** C2970(config)# **mls qos srr-queue input dscp-map queue 1 threshold 2 9 10 11 12 13 14 15** C2970(config)# **mls qos srr-queue input dscp-map queue 1 threshold 3 0 1 2 3 4 5 6 7** C2970(config)# **mls qos srr-queue input dscp-map queue 1 threshold 3 32** C2970(config)# **mls qos srr-queue input dscp-map queue 2 threshold 1 16 17 18 19 20 21 22 23** C2970(config)# **mls qos srr-queue input dscp-map queue 2 threshold 2 33 34 35 36 37 38 39 48** C2970(config)# **mls qos srr-queue input dscp-map queue 2 threshold 2 49 50 51 52 53 54 55 56** C2970(config)# **mls qos srr-queue input dscp-map queue 2 threshold 2 57 58 59 60 61 62 63** C2970(config)# **mls qos srr-queue input dscp-map queue 2 threshold 3 24 25 26 27 28 29 30 31** C2970(config)# **mls qos srr-queue input dscp-map queue 2 threshold 3 40 41 42 43 44 45 46 47**

Table 2-3 Catalyst 2970/3560/3750 Auto-QoS Generated Configuration

The switch automatically maps DSCP values to an egress queue and to a threshold ID.	C2970(config)# **no mls qos srr-queue output dscp-map** C2970(config)# **mls qos srr-queue output dscp-map queue 1 threshold 3 40 41 42 43 44 45 46 47** C2970(config)# **mls qos srr-queue output dscp-map queue 2 threshold 3 24 25 26 27 28 29 30 31** C2970(config)# **mls qos srr-queue output dscp-map queue 2 threshold 3 48 49 50 51 52 53 54 55** C2970(config)# **mls qos srr-queue output dscp-map queue 2 threshold 3 56 57 58 59 60 61 62 63** C2970(config)# **mls qos srr-queue output dscp-map queue 3 threshold 3 16 17 18 19 20 21 22 23** C2970(config)# **mls qos srr-queue output dscp-map queue 3 threshold 3 32 33 34 35 36 37 38 39** C2970(config)# **mls qos srr-queue output dscp-map queue 4 threshold 1 8** C2970(config)# **mls qos srr-queue output dscp-map queue 4 threshold 2 9 10 11 12 13 14 15** C2970(config)# **mls qos srr-queue output dscp-map queue 4 threshold 3 0 1 2 3 4 5 6 7**
The switch automatically sets up the ingress queues, with queue 2 as the priority queue and queue 1 in shared mode. The switch also configures the bandwidth and buffer size for the ingress queues.	C2970(config)# **no mls qos srr-queue input priority-queue 1** C2970(config)# **no mls qos srr-queue input priority-queue 2** C2970(config)# **mls qos srr-queue input bandwidth 90 10** C2970(config)# **mls qos srr-queue input threshold 1 8 16** C2970(config)# **mls qos srr-queue input threshold 2 34 66** C2970(config)# **mls qos srr-queue input buffers 67 33**
The switch automatically configures the egress queue buffer sizes. It configures the bandwidth and the SRR mode (shaped or shared) on the egress queues mapped to the port.	C2970(config)# **mls qos queue-set output 1 threshold 1 138 138 92 138** C2970(config)# **mls qos queue-set output 1 threshold 2 138 138 92 400** C2970(config)# **mls qos queue-set output 1 threshold 3 36 77 100 318** C2970(config)# **mls qos queue-set output 1 threshold 4 20 50 67 400** C2970(config)# **mls qos queue-set output 2 threshold 1 149 149 100 149** C2970(config)# **mls qos queue-set output 2 threshold 2 118 118 100 235** C2970(config)# **mls qos queue-set output 2 threshold 3 41 68 100 272** C2970(config)# **mls qos queue-set output 2 threshold 4 42 72 100 242** C2970(config)# **mls qos queue-set output 1 buffers 10 10 26 54** C2970(config)# **mls qos queue-set output 2 buffers 16 6 17 61** C2970(config-if)# **srr-queue bandwidth shape 10 0 0 0** C2970(config-if)# **srr-queue bandwidth share 10 10 60 20**

Table 2-3 Catalyst 2970/3560/3750 Auto-QoS Generated Configuration

If you entered 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.	C2970(config-if)# **mls qos trust cos** C2970(config-if)# **mls qos trust dscp**
If you entered the **auto qos voip cisco-phone** command, the switch automatically enables the trusted boundary feature, which uses the CDP to detect the presence or absence of a Cisco IP Phone.	C2970(config-if)# **mls qos trust device cisco-phone**
If you entered the **auto qos voip cisco-softphone** command, the switch automatically creates class maps and policy maps.	C2970(config)# **mls qos map policed-dscp 24 26 46 to 0** C2970(config)# **class-map match-all AutoQoS-VoIP-RTP-Trust** C2970(config-cmap)# **match ip dscp ef** C2970(config)# **class-map match-all AutoQoS-VoIP-Control-Trust** C2970(config-cmap)# **match ip dscp cs3 af31** C2970(config)# **policy-map AutoQoS-Police-SoftPhone** C2970(config-pmap)# **class AutoQoS-VoIP-RTP-Trust** C2970(config-pmap-c)# **set dscp ef** C2970(config-pmap-c)# **police 320000 8000 exceed-action policed-dscp-transmit** C2970(config-pmap)# **class AutoQoS-VoIP-Control-Trust** C2970(config-pmap-c)# **set dscp cs3** C2970(config-pmap-c)# **police 32000 8000 exceed-action policed-dscp-transmit**
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.	C2970(config-if)# **service-policy input AutoQoS-Police-SoftPhone**

Catalyst 2970/3560/3750—Untrusted PC + SoftPhone with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

The Untrusted PC + SoftPhone + Scavenger model configuration for the Catalyst 2970/3560/3750 is identical to the Catalyst 3550 for the same access edge model and is shown below.

Example 2-28 Catalyst 2970/3560/3750—Untrusted PC + SoftPhone + Scavenger Model Configuration

```
CAT2970(config)#mls qos map policed-dscp  0 24 46 to 8
    ! Excess traffic marked 0 or CS3 or EF will be remarked to CS1
CAT2970(config)#
CAT2970(config)#class-map match-all SOFTPHONE-VOICE
CAT2970(config-cmap)#  match access-group name SOFTPHONE-VOICE
CAT2970(config-cmap)#class-map match-all SOFTPHONE-SIGNALING
CAT2970(config-cmap)#  match access-group name SOFTPHONE-SIGNALING
CAT2970(config-cmap)#exit
CAT2970(config)#
CAT2970(config)#policy-map SOFTPHONE-PC
CAT2970(config-pmap)#class SOFTPHONE-VOICE
CAT2970(config-pmap-c)# set ip dscp 46          ! Softphone VoIP is marked to DSCP EF
CAT2970(config-pmap-c)# police 128000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile SoftPhone voice traffic is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class SOFTPHONE-SIGNALING
CAT2970(config-pmap-c)# set ip dscp 24              ! Signaling is marked to DSCP CS3
CAT2970(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile Signaling traffic is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class class-default
CAT2970(config-pmap-c)# set ip dscp 0
CAT2970(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)# exit
CAT2970(config-pmap)#exit
CAT2970(config)#
CAT2970(config)#interface GigabitEthernet0/1
CAT2970(config-if)# service-policy input SOFTPHONE-PC           ! Applies policy to int
CAT2970(config-if)#exit
CAT2970(config)#
CAT2970(config)#ip access list  extended SOFTPHONE-VOICE
CAT2970(config-ext-nacl)# permit udp any any range 16384 32767             ! VoIP ports
CAT2970(config-ext-nacl)#
CAT2970(config-ext-nacl)#ip access list  extended SOFTPHONE-SIGNALING
CAT2970(config-ext-nacl)# permit tcp any any range 2000 2002               ! SCCP ports
CAT2970(config-ext-nacl)#end
CAT2970#
```

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for the Untrusted PC + SoftPhone model include the following:

- show mls qos
- show mls qos map
- show mls qos interface
- show mls qos interface policers
- show class-map
- show policy-map
- show policy interface

Catalyst 2970/3560/3750—Untrusted Server with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

The Untrusted Multi-Application Server model configuration for the Catalyst 2970/3560/3750 is identical to the Catalyst 3550 and is shown below.

Example 2-29 Catalyst 2970/3560/3750—Untrusted Multi-Application Server with Scavenger-Class QoS Model Configuration

```
CAT2970(config)#mls qos map policed-dscp  0 10 18 25 to 8
  ! Excess traffic marked 0 or AF11 or AF21 or DSCP 25 will be remarked to CS1
CAT2970(config)#
CAT2970(config)#class-map SAP
CAT2970(config-cmap)# match access-group name SAP
CAT2970(config-cmap)#class-map LOTUS
CAT2970(config-cmap)# match access-group name LOTUS
CAT2970(config-cmap)#class-map IMAP
CAT2970(config-cmap)# match access-group name IMAP
CAT2970(config-cmap)#exit
CAT2970(config)#
CAT2970(config)#policy-map UNTRUSTED-SERVER
CAT2970(config-pmap)#class SAP
CAT2970(config-pmap-c)# set ip dscp 25              ! SAP is marked as Mission-Critical
CAT2970(config-pmap-c)# police 15000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile SAP is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class LOTUS
CAT2970(config-pmap-c)# set ip dscp 18              ! Lotus is marked as Transactional
CAT2970(config-pmap-c)# police 35000000 8000 exceed-action policed-dscp-transmit
! Out-of-profile LOTUS is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class IMAP
CAT2970(config-pmap-c)# set ip dscp 10                  ! IMAP is marked as Bulk Data
CAT2970(config-pmap-c)# police 50000000 8000 exceed-action policed-dscp-transmit
! Out-of-profile IMAP is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class class-default
CAT2970(config-pmap-c)# set ip dscp 0
CAT2970(config-pmap-c)# police 1000000 8000 exceed-action policed-dscp-transmit
! Out-of-profile excess data traffic is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)# exit
CAT2970(config-pmap)#exit
CAT2970(config)#
CAT2970(config)#interface GigabitEthernet0/1
CAT2970(config-if)# service-policy input UNTRUSTED-SERVER
CAT2970(config-if)#exit
CAT2970(config)#
CAT2970(config)#ip access list  extended SAP
CAT2970(config-ext-nacl)# permit tcp any range 3200 3203 any
CAT2970(config-ext-nacl)# permit tcp any eq 3600 any
CAT2970(config-ext-nacl)#
CAT2970(config-ext-nacl)#ip access list  extended LOTUS
CAT2970(config-ext-nacl)# permit tcp any eq 1352 any
CAT2970(config-ext-nacl)#
CAT2970(config-ext-nacl)#ip access list  extended IMAP
CAT2970(config-ext-nacl)# permit tcp any eq 143 any
CAT2970(config-ext-nacl)# permit tcp any eq 220 any
CAT2970(config-ext-nacl)#end
CAT2970#
```

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for the Untrusted Multi-Application Server model include the following:

- show mls qos
- show mls qos map
- show mls qos interface
- show mls qos interface policers
- show class-map
- show policy-map
- show policy interface

Catalyst 2970/3560/3750—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

At the time of writing, the Catalyst 2970/3560/3750 does not fully support per-port/per-VLAN policing due to hardware restrictions. Therefore, access lists are required to match voice and signaling traffic sourced from the VVLAN. These ACLs require the administrator to specify the VVLAN subnet information. The configuration for a Conditionally-Trusted IP Phone and PC (Basic Model) for a Catalyst 2970/3560/3750 is shown below.

Example 2-30 Catalyst 2970/3560/3750—Conditionally-Trusted IP Phone + PC + Scavenger (Basic) Model Configuration

```
CAT2970(config)#mls qos map cos-dscp 0 8 16 24 32 46 48 56
    ! Modifies CoS-to-DSCP mapping to map CoS 5 to DSCP EF
CAT2970(config)#mls qos map policed-dscp 0 24 to 8
    ! Excess VVLAN & DVLAN traffic will be remarked to Scavenger (CS1)
CAT2970(config)#
CAT2970(config)#
CAT2970(config)#class-map match-all VVLAN-VOICE
CAT2970(config-cmap)#  match access-group name VVLAN-VOICE
CAT2970(config-cmap)#
CAT2970(config-cmap)#class-map match-all VVLAN-CALL-SIGNALING
CAT2970(config-cmap)#  match access-group name VVLAN-CALL-SIGNALING
CAT2970(config-cmap)#
CAT2970(config-cmap)#class-map match-all VVLAN-ANY
CAT2970(config-cmap)#  match access-group name VVLAN-ANY
CAT2970(config-cmap)#
CAT2970(config-cmap)#
CAT2970(config-cmap)#policy-map IPPHONE+PC-BASIC
CAT2970(config-pmap)#class VVLAN-VOICE
CAT2970(config-pmap-c)# set ip dscp 46                                  ! DSCP EF (Voice)
CAT2970(config-pmap-c)# police 128000 8000 exceed-action drop
```

```
    ! Only one voice call is permitted per switchport VVLAN
CAT2970(config-pmap-c)#class VVLAN-CALL-SIGNALING
CAT2970(config-pmap-c)# set ip dscp 24                        ! DSCP CS3 (Call-Signaling)
CAT2970(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile call signaling is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class VVLAN-ANY
CAT2970(config-pmap-c)# set ip dscp 0
CAT2970(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Unauthorized VVLAN traffic is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class class-default
CAT2970(config-pmap-c)# set ip dscp 0
CAT2970(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)# exit
CAT2970(config-pmap)#exit
CAT2970(config)#
CAT2970(config)#
CAT2970(config)#interface GigabitEthernet0/1
CAT2970(config-if)# switchport access vlan 10                                  ! DVLAN
CAT2970(config-if)# switchport voice vlan 110                                  ! VVLAN
CAT2970(config-if)# service-policy input IPPHONE+PC-BASIC           ! Attaches policy
CAT2970(config-if)#exit
CAT2970(config)#
CAT2970(config)#
CAT2970(config)#ip access list  extended VVLAN-VOICE
CAT2970(config-ext-nacl)#permit udp 10.1.110.0 0.0.0.255
    any range 16384 32767
    ! Voice is matched by VVLAN subnet and VoIP UDP port-range
CAT2970(config-ext-nacl)#exit
CAT2970(config)#
CAT2970(config)#ip access list  extended VVLAN-CALL-SIGNALING
CAT2970(config-ext-nacl)#permit tcp 10.1.110.0 0.0.0.255
any range 2000 2002
    ! Call Signaling is matched by VVLAN subnet and Call-Signaling TCP port-range(s)
CAT2970(config-ext-nacl)#exit
CAT2970(config)#
CAT2970(config)#ip access list  extended VVLAN-ANY
CAT2970(config-ext-nacl)# permit ip 10.1.110.0 0.0.0.255 any
! Matches all other traffic sourced from the VVLAN subnet
CAT2970(config-ext-nacl)#end
CAT2970#
```

Note At the time of writing, the Catalyst 2970/3560/3750 does not support a trust statement (such as **mls qos trust device cisco-phone**) in conjunction with a service-policy input statement applied to given port at the same time. While this may be configurable, if the switch is reset, one or the other statement may be removed when the switch reloads. This limitation is to be addressed; consult the latest Catalyst 2970/3560/3750 QoS documentation for updates on this limitation.

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for the Conditionally-Trusted IP Phone and PC with Scavenger-Class QoS (Basic) model include the following:

- show mls qos
- show mls qos map
- show mls qos interface
- show mls qos interface policers

- show class-map
- show policy-map
- show policy interface

Catalyst 2970/3560/3750—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

Building on the previous model, PC applications such as Interactive Video, Mission-Critical Data, Transactional Data, and Bulk Data are identified by access lists. The configuration for the Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model for the Catalyst 2970/3560/3750 is shown below.

Example 2-31 Catalyst 2970/3560/3750—Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model Configuration

```
CAT2970(config)#mls qos map cos-dscp 0 8 16 24 32 46 48 56
    ! Modifies CoS-to-DSCP mapping to map CoS 5 to DSCP EF
CAT2970(config)#mls qos map policed-dscp 0 10 18 24 25 34 to 8
! Excess DVLAN traffic marked 0, AF11, AF21, CS3, DSCP 25
! and AF41 will be remarked to Scavenger (CS1)
CAT2970(config)#
CAT2970(config)#
CAT2970(config)#class-map match-all VVLAN-VOICE
CAT2970(config-cmap)#  match access-group name VVLAN-VOICE
CAT2970(config-cmap)#
CAT2970(config-cmap)#class-map match-all VVLAN-CALL-SIGNALING
CAT2970(config-cmap)#  match access-group name VVLAN-CALL-SIGNALING
CAT2970(config-cmap)#
CAT2970(config-cmap)#class-map match-all VVLAN-ANY
CAT2970(config-cmap)#  match access-group name VVLAN-ANY
CAT2970(config-cmap)#
CAT2970(config-cmap)#class-map match-all DVLAN-PC-VIDEO
CAT2970(config-cmap)# match access-group name DVLAN-PC-VIDEO
CAT2970(config-cmap)#
CAT2970(config-cmap)#class-map match-all DVLAN-MISSION-CRITICAL-DATA
CAT2970(config-cmap)# match access-group name DVLAN-MISSION-CRITICAL-DATA
CAT2970(config-cmap)#
CAT2970(config-cmap)#class-map match-all DVLAN-TRANSACTIONAL-DATA
CAT2970(config-cmap)# match access-group name DVLAN-TRANSACTIONAL-DATA
CAT2970(config-cmap)#
CAT2970(config-cmap)#class-map match-all DVLAN-BULK-DATA
CAT2970(config-cmap)# match access-group name DVLAN-BULK-DATA
CAT2970(config-cmap)#exit
CAT2970(config)#
CAT2970(config)#policy-map IPPHONE+PC-ADVANCED
CAT2970(config-pmap)#class VVLAN-VOICE
CAT2970(config-pmap-c)# set ip dscp 46                                  ! DSCP EF (Voice)
CAT2970(config-pmap-c)# police 128000 8000 exceed-action drop
    ! Only one voice call is permitted per switchport VVLAN
```

```
CAT2970(config-pmap-c)#class VVLAN-CALL-SIGNALING
CAT2970(config-pmap-c)# set ip dscp 24                        ! DSCP CS3 (Call-Signaling)
CAT2970(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile call signaling is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class VVLAN-ANY
CAT2970(config-pmap-c)# set ip dscp 0
CAT2970(config-pmap-c)# police 32000 8000 exceed-action policed-dscp-transmit
    ! Unauthorized VVLAN traffic is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class DVLAN-PC-VIDEO
CAT2970(config-pmap-c)# set ip dscp 34                    ! DSCP AF41 (Interactive-Video)
CAT2970(config-pmap-c)# police 496000 8000 exceed-action policed-dscp-transmit
    ! Only one IP/VC stream will be permitted per switchport
CAT2970(config-pmap-c)#class DVLAN-MISSION-CRITICAL-DATA
CAT2970(config-pmap-c)# set ip dscp 25                    ! Interim Mission-Critical Data
CAT2970(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile Mission-Critical Data is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class DVLAN-TRANSACTIONAL-DATA
CAT2970(config-pmap-c)# set ip dscp 18                    ! DSCP AF21 (Transactional Data)
CAT2970(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile Transactional Data is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class DVLAN-BULK-DATA
CAT2970(config-pmap-c)# set ip dscp 10                        ! DSCP AF11 (Bulk Data)
CAT2970(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile Bulk Data is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)#class class-default
CAT2970(config-pmap-c)# set ip dscp 0
CAT2970(config-pmap-c)# police 5000000 8000 exceed-action policed-dscp-transmit
    ! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT2970(config-pmap-c)# exit
CAT2970(config-pmap)#exit
CAT2970(config)#
CAT2970(config)#interface GigabitEthernet0/1
CAT2970(config-if)# switchport access vlan 10                                  ! DVLAN
CAT2970(config-if)# switchport voice vlan 110                                  ! VVLAN
CAT2970(config-if)# service-policy input IPPHONE+PC-ADVANCED        ! Attaches Policy
CAT2970(config-if)#exit
CAT2970(config)#
CAT2970(config)#
CAT2970(config)#ip access list  extended VVLAN-VOICE
CAT2970(config-ext-nacl)#permit udp 10.1.110.0 0.0.0.255
    any range 16384 32767
    ! Voice is matched by VVLAN subnet and DSCP EF
CAT2970(config-ext-nacl)#exit
CAT2970(config)#
CAT2970(config)#ip access list  extended VVLAN-CALL-SIGNALING
CAT2970(config-ext-nacl)#permit tcp 10.1.110.0 0.0.0.255
any range 2000 2002
    ! Call Signaling is matched by VVLAN subnet Call-Signaling TCP port-range(s)
CAT2970(config-ext-nacl)#exit
CAT2970(config)#
CAT2970(config)#ip access list  extended VVLAN-ANY
CAT2970(config-ext-nacl)# permit ip 10.1.110.0 0.0.0.255 any
! Matches all other traffic sourced from the VVLAN subnet
CAT2970(config-ext-nacl)#
CAT2970(config-ext-nacl)#ip access list  extended DVLAN-PC-VIDEO
CAT2970(config-ext-nacl)# permit udp any any range 16384 32767                 ! IP/VC
CAT2970(config-ext-nacl)#
CAT2970(config-ext-nacl)#ip access list  extended DVLAN-MISSION-CRITICAL-DATA
CAT2970(config-ext-nacl)# permit tcp any any range 3200 3203                   ! SAP
CAT2970(config-ext-nacl)# permit tcp any any eq 3600                           ! SAP
CAT2970(config-ext-nacl)# permit tcp any any range 2000 2002                   ! SCCP
CAT2970(config-ext-nacl)#
CAT2970(config-ext-nacl)#ip access list  extended DVLAN-TRANSACTIONAL-DATA
CAT2970(config-ext-nacl)# permit tcp any any eq 1352                           ! Lotus
```

```
CAT2970(config-ext-nacl)#
CAT2970(config-ext-nacl)#ip access list  extended DVLAN-BULK-DATA
CAT2970(config-ext-nacl)# permit tcp any any eq 143                          ! IMAP
CAT2970(config-ext-nacl)# permit tcp any any eq 220                          ! IMAP
CAT2970(config-ext-nacl)#end
CAT2970#
```

Note At the time of writing, the Catalyst 2970/3560/3750 does not support a trust statement (such as **mls qos trust device cisco-phone**) in conjunction with a service-policy input statement applied to given port at the same time. While this may be configurable, if the switch is reset, one or the other statement may be removed when the switch reloads. This limitation is to be addressed; consult the latest Catalyst 2970/3560/3750 QoS documentation for updates on this limitation.

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for the Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model include the following:

- show mls qos
- show mls qos map
- show mls qos interface
- show mls qos interface policers
- show class-map
- show policy-map
- show policy interface

Catalyst 2970/3560/3750—Queuing and Dropping

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Configuration

For the most part, the Catalyst 2970/3560/3750 is relatively compatible in QoS features and syntax with the Catalyst 3550, except with respect to queuing and dropping.

The Catalyst 2970/3560/3750 supports four egress queues, which can be configured on a per-interface basis to operate in either 4Q3T or 1P3Q3T modes. Additionally, the Catalyst 2970/3560/3750 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.

However, unlike the Catalyst 2950 and 3550, the Catalyst 2970/3560/3750 has Queue 1 (not Queue 4) as the optional priority queue. In a converged campus environment it is recommended to enable the priority queue via the **priority-queue out** interface command.

Note The Catalyst 2970/3560/3750 also supports two configurable ingress queues (normal and expedite). Ingress scheduling, however, is rarely—if ever—required, as it only becomes enabled if the combined input rates from any/all switch ports exceed the switch fabric's capacity. 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 three remaining egress queues on the Catalyst 2970/3560/3750 are scheduled by a Shaped Round-Robin (SRR) algorithm, which can be configured to operate in shaped mode or in shared mode. In shaped mode, assigned bandwidth is limited to the defined amount; in shared mode, any unused bandwidth is shared among other classes (as needed).

Shaped or Shared bandwidth weights can be assigned to a queue via the **srr-queue bandwidth shape** and **srr-queue bandwidth share** interface commands. Shaped mode weights override shared mode weights and use an inverse ratio (1/weight) to determine the shaping bandwidth for the queue. Furthermore, if shaped weights are set to 0, then the queue is operating in shared mode. For example, the following interface command **srr-queue bandwidth shape 3 0 0 0** would shape Q1 to 1/3 of the available bandwidth and set all other queues to operate in sharing mode.

To make the queuing structure consistent with examples provided for previously discussed platforms, Queues 2 through 4 should be set to operate in shared mode (which is the default mode of operation on Queues 2 through 4). The ratio of the shared weights determines the relative bandwidth allocations (the absolute values are meaningless). The ratio of the shared weights determines the relative bandwidth allocations (the absolute values are meaningless). Since the PQ of the Catalyst 2970/3560/3750 is Q1 (not Q4 as in the Catalyst 3550), the entire queuing model can be flipped upside down, with Q2 representing the Critical Data queue, Q3 representing the Best Effort queue, and Q4 representing the Scavenger/Bulk queue. Therefore, shared weights of 70, 25, and 5 can be assigned to Queues 2, 3, and 4, respectively.

Additionally, the Catalyst 2970/3560/3750 supports three Weighted Tail Drop (WTD) thresholds per queue. Two of these thresholds are configurable (explicit); the third is non-configurable (implicit), as it is set to the queue-full state (100%). These thresholds can be defined with the **mls qos queue-set output** qset-id **threshold** global command. The only queues that these thresholds need defining (away from defaults) are Queues 2 and 4. In Queue 2, it is recommended to set the first threshold to 70% and the second to 80%, which leaves the third (implicit) threshold set at 100% (the tail of the queue). In Queue 4, it is recommended to set the first threshold to 40%, leaving the default values for both the second and third thresholds at 100%.

Once the queues and thresholds have been defined, traffic can be assigned to queues and thresholds either by CoS values or DSCP values, using the **mls qos srr-queue output cos-map queue** and **mls qos srr-queue output dscp-map queue** global commands, respectively. While DSCP-to-Queue/Threshold maps override CoS-to-Queue/Threshold maps, these mappings should be as consistent as possible to ensure predictable behavior and simplify troubleshooting.

That being said, CoS 0/DSCP 0 (Best Effort traffic) should be mapped to Queue 3 Threshold 3 (the tail of the queue), as 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/Threshold mapping assigning DSCP CS1 to Queue 4 Threshold 1 (previously set at 40%); 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%).

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%). This limits Network Management and Transactional Data to a subset of the 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%). This allows for preferential treatment of call signaling traffic within Queue 2.

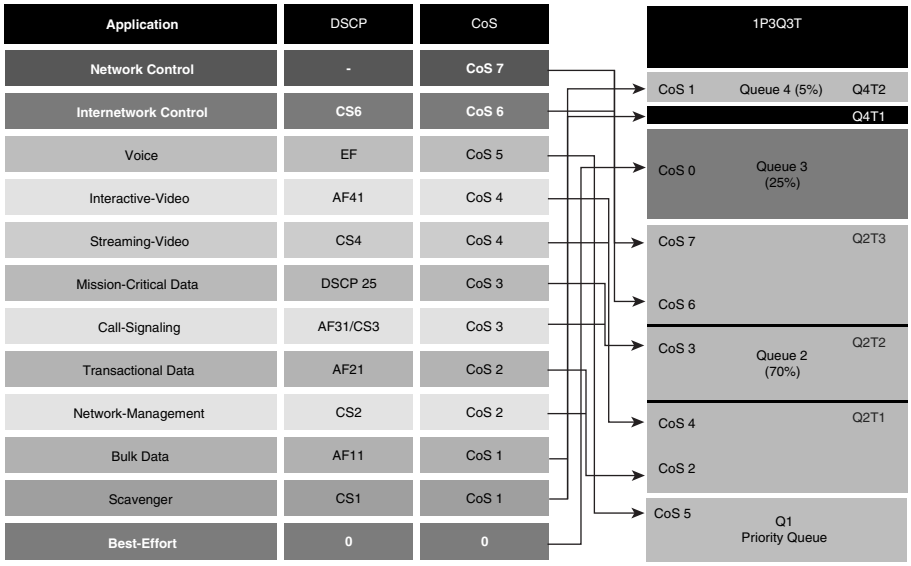
CoS 4 and DSCP CS4, AF41, AF42, and AF43 (decimal values 32, 34, 36, and 38, respectively) can be assigned to Queue 2 Threshold 1. In this manner, video (both Interactive and Streaming) does not drown out call signaling or Network/Internetwork Control traffic within Queue 2.

CoS 5 and DSCP EF (decimal value 46) should be assigned to Queue 1 Threshold 3, as 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.

These recommended Catalyst 2970/3560/3750 CoS/DSCP to Queue/Threshold assignments are illustrated in Figure 2-18.

Figure 2-18 Catalyst 2970/3560/3750 1P3Q3T Queuing Model



The Catalyst 2970/3560/3750 queuing and dropping configuration recommendations are shown below.

Example 2-32 Catalyst 2970/3560/3750—Queuing and Dropping

```
CAT2970(config)#mls qos srr-queue output cos-map queue 1 threshold 3  5
    ! Maps CoS 5 to Queue 1 Threshold 3 (Voice gets all of Queue 1)
CAT2970(config)#mls qos srr-queue output cos-map queue 2 threshold 1  2 4
    ! Maps CoS 2 and CoS 4 to Queue 2 Threshold 1
CAT2970(config)#mls qos srr-queue output cos-map queue 2 threshold 2  3
    ! Maps CoS 3 to Queue 2 Threshold 2
CAT2970(config)#mls qos srr-queue output cos-map queue 2 threshold 3  6 7
    ! Maps CoS 6 and CoS 7 to Queue 2 Threshold 3
CAT2970(config)#mls qos srr-queue output cos-map queue 3 threshold 3  0
    ! Maps CoS 0 to Queue 3 Threshold 3 (Best Efforts gets all of Q3)
```

```
CAT2970(config)#mls qos srr-queue output cos-map queue 4 threshold 3  1
    ! Maps CoS1 to Queue 4 Threshold 3 (Scavenger/Bulk gets all of Q4)
CAT2970(config)#
CAT2970(config)#
CAT2970(config)#mls qos srr-queue output dscp-map queue 1 threshold 3  46
    ! Maps DSCP EF (Voice) to Queue 1 Threshold 3
CAT2970(config)#mls qos srr-queue output dscp-map queue 2 threshold 1  16
    ! Maps DSCP CS2 (Network Management) to Queue 2 Threshold 1
CAT2970(config)#mls qos srr-queue output dscp-map queue 2 threshold 1  18 20 22
    ! Maps DSCP AF21, AF22, AF23 (Transactional Data) to Queue 2 Threshold 1
CAT2970(config)#mls qos srr-queue output dscp-map queue 2 threshold 1  25
    ! Maps DSCP 25 (Mission-Critical Data) to Queue 2 Threshold 1
CAT2970(config)#mls qos srr-queue output dscp-map queue 2 threshold 1  32
    ! Maps DSCP CS4 (Streaming Video) to Queue 2 Threshold 1
CAT2970(config)#mls qos srr-queue output dscp-map queue 2 threshold 1  34 36 38
    ! Maps DSCP AF41, AF42, AF43 (Interactive-Video) to Queue 2 Threshold 1
CAT2970(config)#mls qos srr-queue output dscp-map queue 2 threshold 2  24 26
    ! Maps DSCP CS3 and DSCP AF31 (Call-Signaling) to Queue 2 Threshold 2
CAT2970(config)#mls qos srr-queue output dscp-map queue 2 threshold 3  48 56
    ! Maps DSCP CS6 and CS7 (Network/Internetwork) to Queue 2 Threshold 3
CAT2970(config)#mls qos srr-queue output dscp-map queue 3 threshold 3  0
    ! Maps DSCP 0 (Best Effort) to Queue 3 Threshold 3
CAT2970(config)#mls qos srr-queue output dscp-map queue 4 threshold 1  8
    ! Maps DSCP CS1 (Scavenger) to Queue 4 Threshold 1
CAT2970(config)#mls qos srr-queue output dscp-map queue 4 threshold 3  10 12 14
    ! Maps DSCP AF11, AF12, AF13 (Bulk Data) to Queue 4 Threshold 3
CAT2970(config)#
CAT2970(config)#
CAT2970(config)#mls qos queue-set output 1 threshold 2 70 80 100 100
    ! Sets Q2 Threshold 1 to 70% and Q2 Threshold 2 to 80%
CAT2970(config)#mls qos queue-set output 1 threshold 4 40 100 100 100
    ! Sets Q4 Threshold 1 to 40% and Q4 Threshold 2 to 100%
CAT2970(config)#
CAT2970(config)#interface range GigabitEthernet0/1 - 28
CAT2970(config-if-range)# queue-set 1
    ! Assigns interface to Queue-Set 1 (default)
CAT2970(config-if-range)# srr-queue bandwidth share 1 70 25 5
    ! Q2 gets 70% of remaining BW; Q3 gets 25% and Q4 gets 5%
CAT2970(config-if-range)# srr-queue bandwidth shape  3 0  0  0
    ! Q1 is limited to 1/3 of the total available BW
CAT2970(config-if-range)# priority-queue out
    ! Q1 is enabled as a PQ
CAT2970(config-if-range)#end
CAT2970#
```

Catalyst MLS QoS Verification Commands

Catalyst MLS QoS verification commands for queuing and dropping include the following:

- show mls qos interface buffers
- show mls qos interface queueing
- show mls qos queue-set
- show mls qos maps cos-output-q
- show mls qos maps dscp-output-q

show mls qos queue-set

The **show mls qos queue-set** verification command returns the configured buffer allocations and defined thresholds for each queue-set.

In the example below, each queue has a default buffer allocation of 25%. Additionally, all WTD Thresholds are set to 100% (the tail of the queue), except for Queue 2 Threshold 1 (set to 70%), Queue 2 Threshold 2 (set to 80%), and Queue 4 Threshold 1 (set to 40%).

Example 2-33 Show MLS QoS Queue-Set Verification for a Catalyst 2970/3560/3750 Switch

```
CAT2970#show mls qos queue-set 1
Queueset: 1
Queue     :       1       2       3       4
----------------------------------------------
buffers   :      25      25      25      25
threshold1:     100      70     100      40
threshold2:     100      80     100     100
reserved  :      50     100      50     100
maximum   :     400     100     400     100
CAT2970#
```

show mls qos maps cos-output-q

The **show mls qos maps cos-output-q** verification command truncates the **show mls qos maps** output to only report the CoS-to-Queue/Threshold mappings for egress queues.

In the example below, CoS 0 is mapped to Q3T3, CoS 1 is mapped to Q4T3, CoS 2 is mapped to Q2T1, CoS 3 is mapped to Q2T2, CoS 4 is mapped to Q2T1, CoS 5 is mapped to Q1T3 (the PQ), and CoS 6 and CoS 7 are mapped to Q2T3.

Example 2-34 Show MLS QoS Maps CoS-Output-Q Verification for a Catalyst 2970/3560/3750 Switch

```
CAT2970#show mls qos maps cos-output-q
   Cos-outputq-threshold map:
               cos:  0   1   2   3   4   5   6   7
               ------------------------------------
  queue-threshold: 3-3 4-3 2-1 2-2 2-1 1-3 2-3 2-3


CAT2970#
```

show mls qos maps dscp-output-q

The **show mls qos maps dscp-output-q** verification command truncates the **show mls qos maps** output to only report the DSCP-to-Queue/Threshold mappings for egress queues. The output is shown in tabular form, with the first digit of the decimal DSCP value in rows and the second digit in columns.

In the example below, only standard DSCP PHBs are being mapped away from the default settings (with the exception of the temporary marking of DSCP 25 for Mission-Critical Data). The other non-standard values may be mapped to reflect the CoS-to-Queue mappings, but for example simplicity this has not been done in this case.

Specifically, DSCP 0 is mapped to Q3T3; DSCP CS1 (8) is mapped to Q4T1; DSCP AF11, AF12, and AF13 (10, 12, 14) are mapped to Q4T3; DSCP CS2 (16) is mapped to Q2T1 as are DSCP AF21, AF22, and AF23 (18, 20, 22); DSCP CS3 (24) and AF31 (26) are mapped to Q2T2; DSCP CS4 (32) is mapped to Q2T1 as are DSCP AF41, AF42, and AF43 (34, 36, 38); DSCP EF (46) is mapped to Q1T3 (the PQ); DSCP CS6 (48) and CS7 (56) are mapped to Q2T3. The non-standard DSCP 25 is mapped to Q2T1.

Example 2-35 Show MLS QoS Maps DSCP-Output-Q Verification for a Catalyst 2970/3560/3750 Switch

```
CAT2970#show mls qos maps dscp-output-q
```

```
Dscp-outputq-threshold map:
  d1 :d2    0     1     2     3     4     5     6     7     8     9
  ------------------------------------------------------------
   0 :    03-03 02-01 02-01 02-01 02-01 02-01 02-01 02-01 04-01 02-01
   1 :    04-03 02-01 04-03 02-01 04-03 02-01 02-01 03-01 02-01 03-01
   2 :    02-01 03-01 02-01 03-01 02-02 02-01 02-02 03-01 03-01 03-01
   3 :    03-01 03-01 02-01 04-01 02-01 04-01 02-01 04-01 02-01 04-01
   4 :    01-01 01-01 01-01 01-01 01-01 01-01 01-03 01-01 02-03 04-01
   5 :    04-01 04-01 04-01 04-01 04-01 04-01 02-03 04-01 04-01 04-01
   6 :    04-01 04-01 04-01 04-01

CAT2970#
```

Catalyst 4500 Supervisor II+/III/IV/V—QoS Considerations and Design

This section includes the following topics:

- Catalyst 4500—Trusted Endpoint Model
- Catalyst 4500—Auto QoS VoIP Model
- Catalyst 4500—Untrusted PC + SoftPhone with Scavenger-Class QoS Model
- Catalyst 4500—Untrusted Server with Scavenger-Class QoS Model
- Catalyst 4500 —Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model
- Catalyst 4500—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model
- Catalyst 4500—Queuing

The Catalyst 4500 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 options for access layer Catalyst 4500 design are shown in Figure 2-19; the distribution and/or core layer recommendations are shown in Figure 2-20.

Figure 2-19 Access Layer Catalyst 4500 QoS Design Options

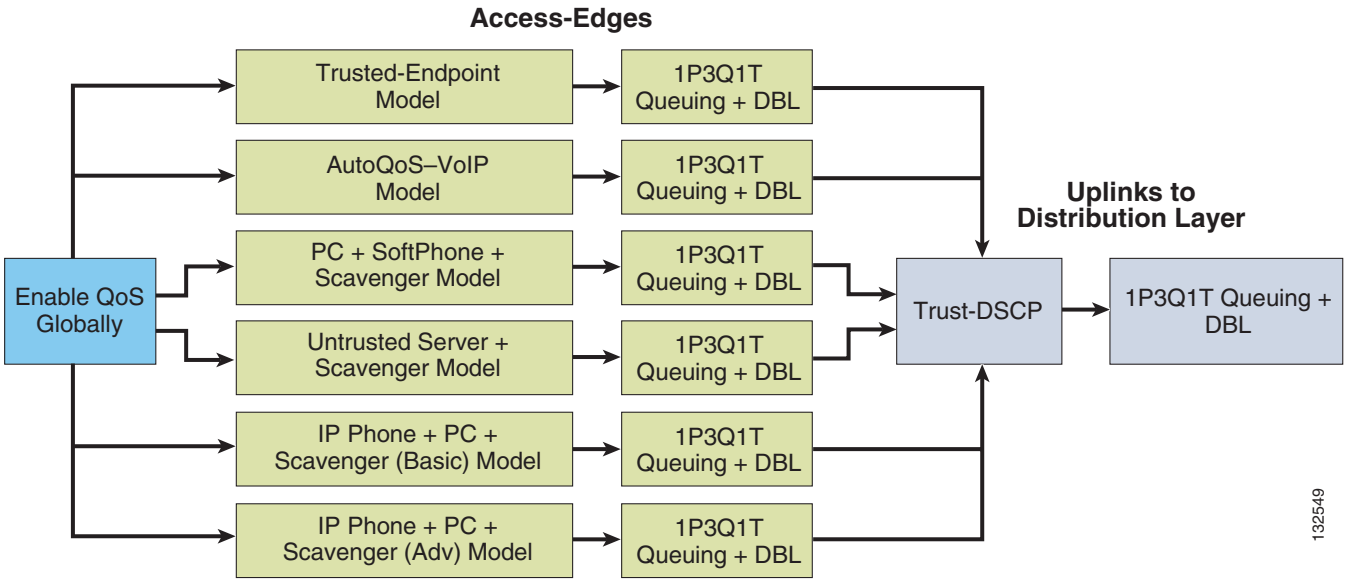
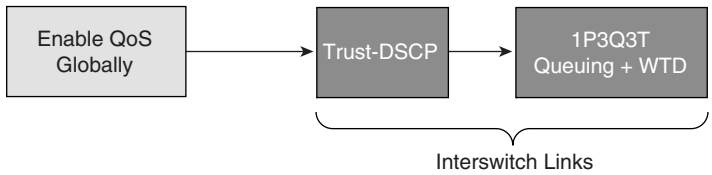


Figure 2-20 Distribution and/or Core Layer Catalyst 4500 QoS Design



Note To narrow the scope of our discussion to the most current and relevant versions of the Catalyst 4500 switch family, only the Catalyst 4500 with Supervisors II+, III, IV and V are examined in this design chapter. For discussions of older versions of Catalyst 4000/4500s, refer to the Cisco Press book, *Cisco Catalyst QoS: Quality of Service in Campus Networks*, by Mike Flannagan, Richard Froom, and Kevin Turek.

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. For example, as with the Catalyst 3550 and 2970/3560/3750, QoS is globally disabled on the Catalyst 4500 by default. However, the command to enable QoS globally on a Catalyst 4500 is simply **qos**, not **mls qos**.

The verification commands are issued in the same manner, with the **mls** keyword omitted. Generally speaking, **show mls qos [...]** verification commands from other Catalyst platforms are translated to **show qos [...]** verification commands on the Catalyst 4500 platforms.

While QoS is globally disabled on the Catalyst 4500, all frames/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, then all DSCP and CoS values are (by default) set to 0 (which is equivalent to an untrusted state on all ports).

The verification command to check whether QoS has been globally enabled on the Catalyst 4500, along with the configuration command, are shown below.

Example 2-36 Enabling QoS Globally on the Catalyst 4500

```
CAT4500#show qos
QoS is disabled globally
IP header DSCP rewrite is enabled

CAT4500#conf term
Enter configuration commands, one per line.  End with CNTL/Z.
CAT4500(config)#qos
CAT4500(config)#end
CAT4500#

CAT4500#show qos
QoS is enabled globally
IP header DSCP rewrite is enabled

CAT4500#
```

Catalyst 4500—Trusted Endpoint Model

This section includes the following topics:

- Configuration
- Catalyst 4500 QoS Verification Commands

Configuration

To enable a given Catalyst 4500 interface to trust the DSCP markings of an endpoint, the **qos trust dscp** interface command is used as shown below.

Example 2-37 Catalyst 4500—Trusted Endpoint Model Configuration

```
CAT4500(config)#interface FastEthernet2/1
CAT4500(config-if)# qos trust dscp
CAT4500(config-if)#end
CAT4500#
```

Catalyst 4500 QoS Verification Commands

Catalyst 4500 QoS verification commands for the Catalyst 4500 Untrusted Endpoint model include the following:

- show qos
- show qos interface

Note Since most Catalyst 4500 verification commands are reasonably similar to the MLS QoS verification commands previously discussed (albeit without the **mls** keyword), to minimize redundancy MLS QoS verification commands that have already been detailed are not repeated in this section.

Catalyst 4500—Auto QoS VoIP Model

The Catalyst 4500 supports AutoQoS VoIP with the following keyword options:

- **auto qos voip cisco-phone**
- **auto qos voip trust**

When you enable AutoQoS VoIP on the Catalyst 4500 by using the **auto qos voip cisco-phone** or the **auto qos voip trust** interface configuration commands, the switch automatically generates a QoS configuration based on the traffic type and ingress packet label and applies the commands listed in Table 2-4 to the interface.

Table 2-4 Catalyst 4500 AutoQoS Generated Configuration

Description	Automatically Generated Command
The switch automatically enables standard QoS and DBL configures the cos-to-DSCP map (maps CoS values in incoming packets to a DSCP value).	C4500(config)# **qos** C4500(config)# **qos map cos 3 to 26** C4500(config)# **qos dbl** C4500(config)# **qos map cos 5 to 46**
The switch automatically configures the DSCP-to-Tx-queue mapping.	C4500(config)# **qos map dscp 24 25 26 27 b28 29 30 31 to tx-queue 4** C4500(config)# **qos map dscp 32 33 34 35 36 37 38 39 to tx-queue 4**
The switch automatically sets the ingress classification on the interface to trust the CoS/DSCP value received in the packet.	C4500(config-if)# **qos trust cos** or C4500(config-if)# **qos trust dscp**
The switch automatically creates a QoS service policy, enables DBL on the policy, and attaches it to the interface.	C4500(config)# **policy-map autoqos-voip-policy** C4500(config-pmap)# **class class-default** C4500(config-pmap-c)# **dbl**
If you entered the **auto qos voip cisco-phone** command, the switch automatically enables the trusted boundary feature, which uses the CDP to detect the presence or absence of a Cisco IP phone.	C4500(config-if)# **qos trust device cisco-phone**
The switch assigns a higher priority for queue 3. Limit for shaping on queue 3 is selected so that it is 33 percent of the link speed. Configure shaping as 33 percent on those ports where sharing is supported. This procedure ensures that the higher-priority queue does not starve other queues.	C4500(config-if)# **tx-queue 3** C4500(config-if-tx-queue)# **priority high** C4500(config-if-tx-queue)# **shape percent 33** C4500(config-if-tx-queue)# **bandwidth percent 33**

Catalyst 4500—Untrusted PC + SoftPhone with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst 4500 QoS Verification Commands

Configuration

The Untrusted PC + Softphone + Scavenger access edge model for Catalyst 4500s is similar to the examples given on previously discussed platforms. A few distinctions exist, such as the absence of the **mls** keyword in defining the policed-DSCP map (along with some slight syntax variation for this command) and the (optional) use of the **kbps** and **mbps** (denoting kilobits and megabits, respectively) within the policing statements.

Example 2-38 Catalyst 4500—Untrusted PC + SoftPhone + Scavenger Model Configuration

```
CAT4500-SUP4(config)#qos map dscp policed 0 24 46 to dscp 8
! Excess traffic marked 0 or CS3 or EF will be remarked to CS1
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#class-map match-all SOFTPHONE-SIGNALING
CAT4500-SUP4(config-cmap)#  match access-group name SOFTPHONE-SIGNALING
CAT4500-SUP4(config-cmap)#class-map match-all SOFTPHONE-VOICE
CAT4500-SUP4(config-cmap)#  match access-group name SOFTPHONE-VOICE
CAT4500-SUP4(config-cmap)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#policy-map SOFTPHONE-PC
CAT4500-SUP4(config-pmap)# class SOFTPHONE-VOICE
CAT4500-SUP4(config-pmap-c)# set ip dscp ef
! Softphone VoIP is marked to DSCP EF
CAT4500-SUP4(config-pmap-c)# police 128 kbps 8000 byte exceed-action
 policed-dscp-transmit
! Out-of-profile SoftPhone voice traffic is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class SOFTPHONE-SIGNALING
CAT4500-SUP4(config-pmap-c)# set ip dscp cs3
    ! SoftPhone call signaling is marked to DSCP CS3
CAT4500-SUP4(config-pmap-c)# police 32 kbps 8000 byte exceed-action
 policed-dscp-transmit
! Out-of-profile Signaling traffic is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class class-default
CAT4500-SUP4(config-pmap-c)# set ip dscp default
CAT4500-SUP4(config-pmap-c)# police 5 mbps 8000 byte exceed-action
 policed-dscp-transmit
! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#interface FastEthernet2/1
CAT4500-SUP4(config-if)# service-policy input SOFTPHONE-PC             ! Applies policy
CAT4500-SUP4(config-if)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#ip access list  extended SOFTPHONE-VOICE
CAT4500-SUP4(config-ext-nacl)# permit udp any any range 16384 32767            ! VoIP
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended SOFTPHONE-SIGNALING
CAT4500-SUP4(config-ext-nacl)# permit tcp any any range 2000 2002              ! SCCP
CAT4500-SUP4(config-ext-nacl)#end
CAT4500-SUP4#
```

Catalyst 4500 QoS Verification Commands

Catalyst 4500 QoS verification commands for the Catalyst 4500 Untrusted PC + Softphone + Scavenger model include the following:

- show qos
- show qos maps
- show qos interface

- show class-map
- show policy-map
- show policy interface

Catalyst 4500—Untrusted Server with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst 4500 QoS Verification Commands

Configuration

The Catalyst 4500 Untrusted Multi-Application Server with Scavenger-Class QoS model is show below. The main changes for the Catalyst 4500 for this model are the syntax defining the policed-DSCP map and the policer definitions (using the abbreviation **mbps** for megabits per second).

Example 2-39 Catalyst 4500—Untrusted Multi-Application Server with Scavenger-Class QoS Model Configuration

```
CAT4500-SUP4(config)#qos map dscp policed 0 10 18 25 to dscp 8
  ! Excess traffic marked 0 or AF11 or AF21 or DSCP 25 will be remarked to CS1
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#class-map SAP
CAT4500-SUP4(config-cmap)# match access-group name SAP
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map LOTUS
CAT4500-SUP4(config-cmap)# match access-group name LOTUS
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map IMAP
CAT4500-SUP4(config-cmap)# match access-group name IMAP
CAT4500-SUP4(config-cmap)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#policy-map UNTRUSTED-SERVER
CAT4500-SUP4(config-pmap)#class SAP
CAT4500-SUP4(config-pmap-c)# set ip dscp 25
! SAP is marked as Mission-Critical (DSCP 25)
CAT4500-SUP4(config-pmap-c)#  police 15 mbps 8000 byte exceed-action
 policed-dscp-transmit
! Out-of-profile SAP is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class LOTUS
CAT4500-SUP4(config-pmap-c)# set ip dscp 18
    ! Lotus is marked as Transactional Data (DSCP AF21)
CAT4500-SUP4(config-pmap-c)#  police 35 mbps 8000 byte exceed-action
 policed-dscp-transmit
! Out-of-profile LOTUS is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class IMAP
CAT4500-SUP4(config-pmap-c)# set ip dscp 10
! IMAP is marked as Bulk Data (DSCP AF11)
CAT4500-SUP4(config-pmap-c)#  police 50 mbps 8000 byte exceed-action
 policed-dscp-transmit
! Out-of-profile IMAP is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class class-default
CAT4500-SUP4(config-pmap-c)# set ip dscp 0
CAT4500-SUP4(config-pmap-c)#  police 1 mbps 8000 byte exceed-action
 policed-dscp-transmit
! Out-of-profile excess data traffic is marked down to Scavenger (CS1)
```

```
CAT4500-SUP4(config-pmap-c)# exit
CAT4500-SUP4(config-pmap)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#interface FastEthernet2/1
CAT4500-SUP4(config-if)# service-policy input UNTRUSTED-SERVER
CAT4500-SUP4(config-if)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#ip access list  extended SAP
CAT4500-SUP4(config-ext-nacl)# permit tcp any range 3200 3203 any
CAT4500-SUP4(config-ext-nacl)# permit tcp any eq 3600 any
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended LOTUS
CAT4500-SUP4(config-ext-nacl)# permit tcp any eq 1352 any
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended IMAP
CAT4500-SUP4(config-ext-nacl)# permit tcp any eq 143 any
CAT4500-SUP4(config-ext-nacl)# permit tcp any eq 220 any
CAT4500-SUP4(config-ext-nacl)#end
CAT4500-SUP4#
```

Catalyst 4500 QoS Verification Commands

Catalyst 4500 QoS verification commands for the Catalyst 4500 Untrusted Server with Scavenger-Class QoS model include the following:

- show qos
- show qos maps
- show qos interface
- show class-map
- show policy-map
- show policy interface

Catalyst 4500 —Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model

This section includes the following topics:

- Configuration
- Catalyst 4500 QoS Verification Commands

Configuration

The Catalyst 4500 does not currently support per-port/per-VLAN policing. Therefore, access lists that include the VVLAN subnet are required to achieve granular policing of the VVLAN and DVLAN subnets, as shown below.

Example 2-40 Catalyst 4500—Conditionally-Trusted IP Phone + PC + Scavenger (Basic) Model Configuration

```
CAT4500-SUP4(config)#qos map cos 5 to dscp 46
! Modifies CoS-to-DSCP mapping to map CoS 5 to DSCP EF
```

```
CAT4500-SUP4(config)#qos map dscp policed 0 24 to dscp 8
    ! Excess DVLAN & VVLAN traffic will be marked down to Scavenger (CS1)
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#class-map match-all VVLAN-VOICE
CAT4500-SUP4(config-cmap)#  match access-group name VVLAN-VOICE
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map match-all VVLAN-CALL-SIGNALING
CAT4500-SUP4(config-cmap)#  match access-group name VVLAN-CALL-SIGNALING
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map match-all VVLAN-ANY
CAT4500-SUP4(config-cmap)#  match access-group name VVLAN-ANY
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#policy-map IPPHONE+PC-BASIC
CAT4500-SUP4(config-pmap)#class VVLAN-VOICE
CAT4500-SUP4(config-pmap-c)# set ip dscp 46                          ! DSCP EF (Voice)
CAT4500-SUP4(config-pmap-c)# police 128 kbps 8000 byte exceed-action drop
    ! Only one voice call is permitted per switchport VVLAN
CAT4500-SUP4(config-pmap-c)#class VVLAN-CALL-SIGNALING
CAT4500-SUP4(config-pmap-c)# set ip dscp 24                ! DSCP CS3 (Call-Signaling)
CAT4500-SUP4(config-pmap-c)# police 32 kbps 8000 byte exceed-action
 policed-dscp-transmit
    ! Out-of-profile call signaling is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class VVLAN-ANY
CAT4500-SUP4(config-pmap-c)# set ip dscp 0
CAT4500-SUP4(config-pmap-c)# police 32 kbps 8000 byte exceed-action
 policed-dscp-transmit
    ! Unauthorized VVLAN traffic is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class class-default
CAT4500-SUP4(config-pmap-c)# set ip dscp 0
CAT4500-SUP4(config-pmap-c)# police 5 mbps 8000 byte exceed-action
 policed-dscp-transmit
! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)# exit
CAT4500-SUP4(config-pmap)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#interface FastEthernet2/1
CAT4500-SUP4(config-if)# switchport access vlan 10                          ! DVLAN
CAT4500-SUP4(config-if)# switchport voice vlan 110                          ! VVLAN
CAT4500-SUP4(config-if)# qos trust device cisco-phone              ! Conditional Trust
CAT4500-SUP4(config-if)# service-policy input IPPHONE+PC-BASIC
CAT4500-SUP4(config-if)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#ip access list  extended VVLAN-VOICE
CAT4500-SUP4(config-ext-nacl)# permit udp 10.1.110.0 0.0.0.255 any
 range 16384 32767
    ! Voice is matched by VVLAN subnet and UDP port-range
CAT4500-SUP4(config-ext-nacl)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#ip access list  extended VVLAN-CALL-SIGNALING
CAT4500-SUP4(config-ext-nacl)# permit tcp 10.1.110.0 0.0.0.255 any
range 2000 2002
    ! Call Signaling is matched by VVLAN subnet and TCP port-range
CAT4500-SUP4(config-ext-nacl)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#ip access list  extended VVLAN-ANY
CAT4500-SUP4(config-ext-nacl)# permit ip 10.1.110.0 0.0.0.255 any
! Matches all other traffic sourced from the VVLAN subnet
CAT4500-SUP4(config-ext-nacl)#end
CAT4500-SUP4#
```

Catalyst 4500 QoS Verification Commands

Catalyst 4500 QoS verification commands for the Conditionally-Trusted IP Phone + PC + Scavenger (Basic) model include the following:

- show qos
- show qos maps
- show qos interface
- show class-map
- show policy-map
- show policy interface

Catalyst 4500—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model

This section includes the following topics:

- Configuration
- Catalyst 4500 QoS Verification Commands

Configuration

Building on the previous model, PC applications, such as Interactive-Video, Mission-Critical Data, Transactional-Data, and Bulk-Data, are identified by access lists. The configuration for the Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model for the Catalyst 4500 is shown below.

Example 2-41 Catalyst 4500—Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model Configuration

```
CAT4500-SUP4(config)#qos map cos 5 to dscp 46
    ! Modifies CoS-to-DSCP mapping to map CoS 5 to DSCP EF
CAT4500-SUP4(config)#qos map dscp policed 0 10 18 24 25 34 to dscp 8
! Excess DVLAN traffic marked 0, AF11, AF21, CS3, DSCP 25
! and AF41 will be remarked to Scavenger (CS1)
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#class-map match-all VVLAN-VOICE
CAT4500-SUP4(config-cmap)#  match access-group name VVLAN-VOICE
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map match-all VVLAN-CALL-SIGNALING
CAT4500-SUP4(config-cmap)#  match access-group name VVLAN-CALL-SIGNALING
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map match-all VVLAN-ANY
CAT4500-SUP4(config-cmap)#  match access-group name VVLAN-ANY
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map match-all DVLAN-PC-VIDEO
CAT4500-SUP4(config-cmap)# match access-group name DVLAN-PC-VIDEO
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map match-all DVLAN-MISSION-CRITICAL-DATA
CAT4500-SUP4(config-cmap)# match access-group name DVLAN-MISSION-CRITICAL-DATA
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map match-all DVLAN-TRANSACTIONAL-DATA
```

```
CAT4500-SUP4(config-cmap)# match access-group name DVLAN-TRANSACTIONAL-DATA
CAT4500-SUP4(config-cmap)#
CAT4500-SUP4(config-cmap)#class-map match-all DVLAN-BULK-DATA
CAT4500-SUP4(config-cmap)# match access-group name DVLAN-BULK-DATA
CAT4500-SUP4(config-cmap)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#policy-map IPPHONE+PC-ADVANCED
CAT4500-SUP4(config-pmap)#class VVLAN-VOICE
CAT4500-SUP4(config-pmap-c)# set ip dscp 46                              ! DSCP EF (Voice)
CAT4500-SUP4(config-pmap-c)# police 128 kbps 8000 byte exceed-action drop
   ! Only one voice call is permitted per switchport VVLAN
CAT4500-SUP4(config-pmap-c)#class VVLAN-CALL-SIGNALING
CAT4500-SUP4(config-pmap-c)# set ip dscp 24                ! DSCP CS3 (Call-Signaling)
CAT4500-SUP4(config-pmap-c)# police 32 kbps 8000 byte exceed-action
 policed-dscp-transmit
   ! Out-of-profile call signaling is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class VVLAN-ANY
CAT4500-SUP4(config-pmap-c)# set ip dscp 0
CAT4500-SUP4(config-pmap-c)# police 32 kbps 8000 byte exceed-action
 policed-dscp-transmit
! Unauthorized VVLAN traffic is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class DVLAN-PC-VIDEO
CAT4500-SUP4(config-pmap-c)# set ip dscp 34                ! DSCP AF41 (Int-Video)
CAT4500-SUP4(config-pmap-c)# police 500 kbps 8000 byte exceed-action
 policed-dscp-transmit
   ! Only one IP/VC stream will be permitted per switchport
CAT4500-SUP4(config-pmap-c)#class DVLAN-MISSION-CRITICAL-DATA
CAT4500-SUP4(config-pmap-c)# set ip dscp 25                ! Interim Mission-Critical
CAT4500-SUP4(config-pmap-c)# police 5 mbps 8000 byte exceed-action
 policed-dscp-transmit
   ! Out-of-profile Mission-Critical Data is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class DVLAN-TRANSACTIONAL-DATA
CAT4500-SUP4(config-pmap-c)# set ip dscp 18                              ! DSCP AF21
CAT4500-SUP4(config-pmap-c)# police 5 mbps 8000 byte exceed-action
 policed-dscp-transmit
   ! Out-of-profile Transactional Data is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class DVLAN-BULK-DATA
CAT4500-SUP4(config-pmap-c)# set ip dscp 10                              ! DSCP AF11
CAT4500-SUP4(config-pmap-c)# police 5 mbps 8000 byte exceed-action
 policed-dscp-transmit
   ! Out-of-profile Bulk Data is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#class class-default
CAT4500-SUP4(config-pmap-c)# set ip dscp 0
CAT4500-SUP4(config-pmap-c)# police 5 mbps 8000 byte exceed-action
 policed-dscp-transmit
   ! Out-of-profile data traffic is marked down to Scavenger (CS1)
CAT4500-SUP4(config-pmap-c)#exit
CAT4500-SUP4(config-pmap)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#interface FastEthernet2/1
CAT4500-SUP4(config-if)# switchport access vlan 10                          ! DVLAN
CAT4500-SUP4(config-if)# switchport voice vlan 110                          ! VVLAN
CAT4500-SUP4(config-if)# qos trust device cisco-phone              ! Conditional Trust
CAT4500-SUP4(config-if)# service-policy input IPPHONE+PC-ADVANCED
CAT4500-SUP4(config-if)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#ip access list  extended VVLAN-VOICE
CAT4500-SUP4(config-ext-nacl)# permit udp 10.1.110.0 0.0.0.255 any
 range 16384 32767
   ! Voice is matched by VVLAN subnet and UDP port-range
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended VVLAN-CALL-SIGNALING
CAT4500-SUP4(config-ext-nacl)# permit tcp 10.1.110.0 0.0.0.255 any
```

```
 range 2000 2002
    ! Call Signaling is matched by VVLAN subnet and TCP port-range
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended VVLAN-ANY
CAT4500-SUP4(config-ext-nacl)# permit ip 10.1.110.0 0.0.0.255 any
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended DVLAN-PC-VIDEO
CAT4500-SUP4(config-ext-nacl)# permit udp any any range 16384 32767
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended DVLAN-MISSION-CRITICAL-DATA
CAT4500-SUP4(config-ext-nacl)# permit tcp any any range 3200 3203              ! SAP
CAT4500-SUP4(config-ext-nacl)# permit tcp any any eq 3600                      ! SAP
CAT4500-SUP4(config-ext-nacl)# permit tcp any any range 2000 2002              ! SCCP
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended DVLAN-TRANSACTIONAL-DATA
CAT4500-SUP4(config-ext-nacl)# permit tcp any any eq 1352                      ! Lotus
CAT4500-SUP4(config-ext-nacl)#
CAT4500-SUP4(config-ext-nacl)#ip access list  extended DVLAN-BULK-DATA
CAT4500-SUP4(config-ext-nacl)# permit tcp any any eq 143                       ! IMAP
CAT4500-SUP4(config-ext-nacl)# permit tcp any any eq 220                       ! IMAP
CAT4500-SUP4(config-ext-nacl)#end
CAT4500-SUP4#
```

Catalyst 4500 QoS Verification Commands

Catalyst 4500 QoS verification commands for the Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) model include the following:

- show qos
- show qos maps
- show qos interface
- show class-map
- show policy-map
- show policy interface

Catalyst 4500—Queuing

This section includes the following topics:

- Configuration
- Catalyst 4500 QoS Verification Commands

Configuration

The Catalyst 4500 supports four egress queues for scheduling, which may be configured in either 4Q1T or 1P3Q1T modes. The strict-priority queue on the Catalyst 4500 is transmit-queue 3.

While tail-drop or WRED thresholds are not supported on the Catalyst 4500, it 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** policy command. A default DBL policy can be applied to all transmit queues, as is shown in the example below.

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 with the **tx-queue 3** interface command followed by the **priority high** interface transmit-queue sub-command. This queue can be defined to be shaped to a peak limit, such as 30%, to allow bandwidth to be available to non-voice applications. This would be valuable in the event that a trust boundary has been compromised and a DoS/worm attack is saturating voice queues.

Bandwidth allocations can also be assigned to queues (for certain interfaces) using the **tx-queue** interface command 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 linecard
- The 2 1000BASE-X ports on the WS-X4232-GB-RJ linecard
- The first 2 ports on the WS-X4418-GB linecard
- The two 1000BASE-X ports on the WS-X4412-2GB-TX linecard

The Catalyst 4500 does not support CoS-to-Queue mappings, only DSCP-to-Queue mappings. These can be defined with the **qos map dscp to tx-queue** global command.

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% of the link's capacity. Furthermore, Q1 can then be used as the Scavenger/Bulk queue, Q2 as the Best-Effort queu,e and Q4 as the preferential queue.

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

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

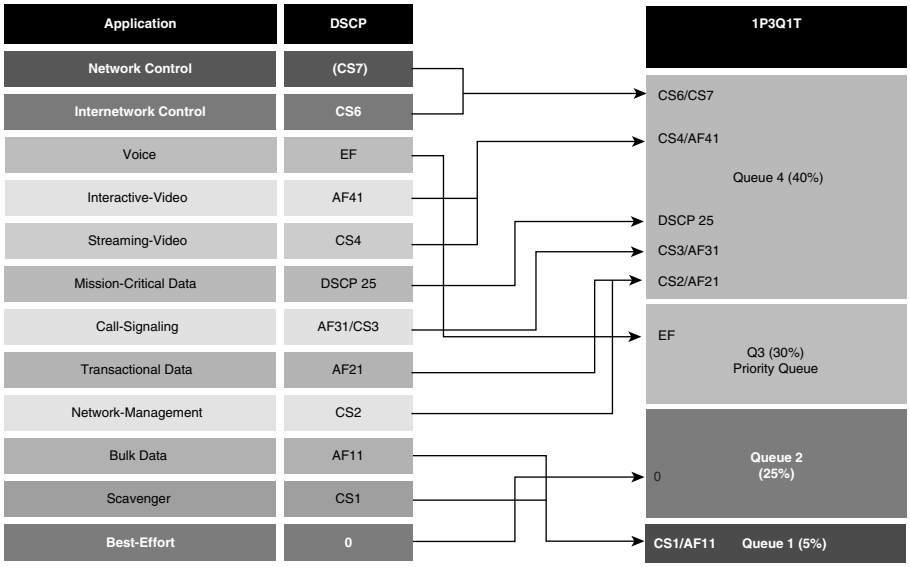
The recommended DSCP-to-Queue assignments for the Catalyst 4500 are as follows:

- DSCP 0 should be assigned to Q2
- DSCP CS1 (Scavenger) and DSCP AF11/AF12/AF13 (Bulk Data) should be assigned to Q1
- DSCP CS2 (Network Management) as well as AF21/AF22/AF23 (Transactional Data) should be assigned to Q4
- DSCP CS3 and AF31 (Call-Signaling) should be assigned to Q4
- DSCP 25 (temporary marking for Mission-Critical Data) should be assigned to Q4
- DSCP CS4 (Streaming Video) and AF41/AF42/AF43 (Interactive Video) should be assigned to Q4
- DSCP EF (Voice) should be assigned to Q3 (the strict priority queue)

- DSCP CS6 (Internetwork Control) and CS7 (Network Control/STP) should be assigned to Q4

The queuing recommendations for the Catalyst 4500 (Supervisors II+, III, IV and V) are shown in Figure 2-21.

Figure 2-21 Catalyst 4500-SupII+/III/IV/V 1P3Q1T Queuing Model



The configurations for enabling queuing on the Catalyst 4500 per these recommendations are shown below. Two separate examples are given, one for a FastEthernet interface that does not support bandwidth allocations and another for a GigabitEthernet interface that does. Some of the DSCP-to-Queue mappings are not required (as they overlap with the default settings), but are shown nonetheless to complete the logic of the example.

Example 2-42 Catalyst 4500—Queuing and Dropping

```
CAT4500-SUP4(config)#qos dbl
    ! Globally enables DBL
CAT4500-SUP4(config)#qos dbl exceed-action ecn
    ! Optional: Enables DBL to mark RFC 3168 ECN bits in the IP ToS Byte
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#qos map dscp 0 to tx-queue 2
    ! Maps DSCP 0 (Best Effort) to Q2
CAT4500-SUP4(config)#qos map dscp 8 10 12 14 to tx-queue 1
    ! Maps DSCP CS1 (Scavenger) and AF11/AF12/AF13 (Bulk) to Q1
CAT4500-SUP4(config)#qos map dscp 16 18 20 22 to tx-queue 4
    ! Maps DSCP CS2 (Net-Mgmt) and AF21/AF22/AF23 (Transactional) to Q4
CAT4500-SUP4(config)#qos map dscp 24 25 26 to tx-queue 4
    ! Maps DSCP CS3 and AF31 (Call-Signaling) and DSCP 25 (MC Data) to Q4
CAT4500-SUP4(config)#qos map dscp 32 34 36 38 to tx-queue 4
    ! Maps DSCP CS4 (Str-Video) and AF41/AF42/AF43 (Int-Video) to Q4
CAT4500-SUP4(config)#qos map dscp 46 to tx-queue 3
    ! Maps DSCP EF (VoIP) to Q3 (PQ)
CAT4500-SUP4(config)#qos map dscp 48 56 to tx-queue 4
    ! Maps DSCP CS6 (Internetwork) and CS7 (Network) Control to Q4
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#policy-map DBL
CAT4500-SUP4(config-pmap)#class class-default
CAT4500-SUP4(config-pmap-c)# dbl                    ! Enables DBL on all traffic flows
CAT4500-SUP4(config-pmap-c)# exit
```

```
CAT4500-SUP4(config-pmap)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#interface range FastEthernet2/1 - 48
CAT4500-SUP4(config-if-range)# service-policy output DBL            ! Applies DBL policy
CAT4500-SUP4(config-if-range)# tx-queue 3
CAT4500-SUP4(config-if-tx-queue)# priority high                     ! Enables Q3 as PQ
CAT4500-SUP4(config-if-tx-queue)# shape percent 30                  ! Shapes PQ to 30%
CAT4500-SUP4(config-if-tx-queue)# exit
CAT4500-SUP4(config-if-range)#exit
CAT4500-SUP4(config)#
CAT4500-SUP4(config)#interface range GigabitEthernet1/1 - 2
CAT4500-SUP4(config-if-range)# service-policy output DBL            ! Applies DBL policy
CAT4500-SUP4(config-if-range)# tx-queue 1
CAT4500-SUP4(config-if-tx-queue)# bandwidth percent 5               ! Q1 gets 5%
CAT4500-SUP4(config-if-tx-queue)# tx-queue 2
CAT4500-SUP4(config-if-tx-queue)# bandwidth percent 25              ! Q2 gets 25%
CAT4500-SUP4(config-if-tx-queue)# tx-queue 3
CAT4500-SUP4(config-if-tx-queue)# priority high                     ! Enables Q3 as PQ
CAT4500-SUP4(config-if-tx-queue)# bandwidth percent 30              ! PQ gets 30%
CAT4500-SUP4(config-if-tx-queue)# shape percent 30                  ! Shapes PQ to 30%
CAT4500-SUP4(config-if-tx-queue)# tx-queue 4
CAT4500-SUP4(config-if-tx-queue)# bandwidth percent 40              ! Q4 gets 40%
CAT4500-SUP4(config-if-tx-queue)#end
CAT4500-SUP4#
```

Catalyst 4500 QoS Verification Commands

Catalyst 4500 QoS verification commands for queuing include the following:

- show qos dbl
- show qos maps dscp tx-queue
- show qos interface

show qos dbl

The Catalyst 4500 **show qos dbl** verification command returns whether or not DBL has been enabled, as well as some of the operating parameters that have been defined for its operation. These parameters include allowing DBL to set RFC 3168 ECN bits in IP headers, as shown in the example below.

Example 2-43 Show QoS DBL Verification for a Catalyst 4500 Switch

```
CAT4500-SUP4#show qos dbl
QOS is enabled globally
DBL is enabled globally
DBL flow includes vlan
DBL flow includes layer4-ports
DBL uses ecn to indicate congestion
DBL exceed-action probability: 15%
DBL max credits: 15
DBL aggressive credit limit: 10
DBL aggressive buffer limit: 2 packets

CAT4500-SUP4#
```

show qos maps dscp tx-queue

The Catalyst 4500 **show qos maps dscp tx-queue** verification command truncates the **show qos maps** output to only report the DSCP-to-Queue mappings for egress queues. The output is shown in tabular form, with the first digit of the decimal DSCP value in rows and the second digit in columns.

In the example below, only DSCP 0 is mapped to Q2; DSCP CS1 (8) and AF11/AF12/AF13 (10/12/14) are mapped to Q1; DSCP CS2 (16) and AF22/AF22/AF23 (18/20/22) are mapped to Q4; DSCP CS3 (24) and AF31 (26) are mapped to Q4, as is the non-standard DSCP 25. DSCP CS4 (32) and AF41/AF42/AF43 (34/36/38) are mapped to Q4, as are DSCP CS6 (48) and CS7 (56). DSCP EF (46) is mapped to Q3.

Example 2-44 Show QoS Maps DSCP Tx-Queue Verification for a Catalyst 4500 Switch

```
CAT4500-SUP4#show qos maps dscp tx-queue
DSCP-TxQueue Mapping Table (dscp = d1d2)
d1 : d2  0  1  2  3  4  5  6  7  8  9
------------------------------------
 0 :    02 01 01 01 01 01 01 01 01 01              ! DSCP 0 => Q2; DSCP CS1 => Q1
 1 :    01 01 01 01 01 01 04 02 04 02              ! DSCP AF11/AF12/AF13 => Q1
                                                   ! DSCP CS2 and AF21 => Q4
 2 :    04 02 04 02 04 04 04 02 02 02              ! DSCP AF22/AF23 => Q4
                                                   ! DSCP CS3, 25 and AF31 => Q4
 3 :    02 02 04 03 04 03 04 03 04 03              ! DSCP CS4 and AF41/AF42/AF43 => Q4
 4 :    03 03 03 03 03 03 03 03 04 04              ! DSCP EF => Q3; DSCP CS6 => Q4
 5 :    04 04 04 04 04 04 04 04 04 04              ! DSCP CS7 => Q4
 6 :    04 04 04 04

CAT4500-SUP4#
```

show qos interface

The Catalyst 4500 **show qos interface** verification command displays the global state of QoS (enabled or not), the trust state of an interface, as well as any queuing/shaping parameters that have been defined for the interface.

In the first example below, the **show qos interface** command is being applied on an access-edge FastEthernet interface that has been configured to conditionally-trust Cisco IP Phones. Furthermore, the output reports that Q3 has been enabled as the priority-queue on this interface and is shaped to 30 Mbps (30%). Bandwidth cannot be assigned for non-priority queues on this interface, as is indicated by the "N/A" entries under the bandwidth column.

In the second example below, the **show qos interface** command is being applied to a GigabitEthernet uplink interface which as been configured to trust-DSCP. As before, Q3 has been enabled as the priority-queue and has been shaped to 30%, which now translates to 300 Mbps. Bandwidth is assignable on this interface and therefore Q1 is allocated 50 Mbps (5%), Q2 is allocated 250 Mbps (25%), Q3 is allocated 300 Mbps (30%), and Q4 is allocated 400 Mbps (40%).

Example 2-45 Show QoS Interface Verification for a Catalyst 4500 Switch

```
CAT4500-SUP4#show qos interface FastEthernet2/1
QoS is enabled globally
Port QoS is enabled
Administrative Port Trust State: 'cos'
Operational Port Trust State: 'cos'
Trust device: cisco-phone
Default DSCP: 0 Default CoS: 0
Appliance trust: none
Tx-Queue    Bandwidth    ShapeRate    Priority    QueueSize
```

```
            (bps)       (bps)                (packets)
  1         N/A         disabled    N/A      240
  2         N/A         disabled    N/A      240
  3         N/A         30000000    high     240
  4         N/A         disabled    N/A      240

CAT4500-SUP4#


CAT4500-SUP4#show qos interface GigabitEthernet1/1
QoS is enabled globally
Port QoS is enabled
Administrative Port Trust State: 'dscp'
Operational Port Trust State: 'dscp'
Trust device: none
Default DSCP: 0 Default CoS: 0
Appliance trust: none
Tx-Queue    Bandwidth    ShapeRate    Priority    QueueSize
            (bps)        (bps)                    (packets)
  1         50000000     disabled     N/A         1920
  2         250000000    disabled     N/A         1920
  3         300000000    300000000    high        1920
  4         700000000    disabled     N/A         1920

CAT4500-SUP4#
```

Catalyst 6500 PFC2/PFC3—QoS Considerations and Design

This section includes the following topics:

- Catalyst 6500 QoS Configuration and Design Overview
- Catalyst 6500—CatOS Defaults and Recommendations
- Catalyst 6500—Trusted Endpoint Model
- Catalyst 6500 Auto QoS VoIP Model
- Catalyst 6500—Untrusted PC + SoftPhone with Scavenger-Class QoS Model
- Catalyst 6500—Untrusted Server with Scavenger-Class QoS Model
- Catalyst 6500—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model
- Catalyst 6500—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model
- Catalyst 6500—Queuing and Dropping
- Catalyst 6500—PFC3 Distribution-Layer (IOS) Per-User Microflow Policing

Catalyst 6500 QoS Configuration and Design Overview

The Catalyst 6500 is the undisputed flagship of the Cisco family of LAN switches, as it is the most powerful and flexible switching platform. As such, it can be found in all three layers of a campus network (Access, Distribution, and Core).

When configured as an access layer switch, the recommended software for the Supervisor is CatOS; when configured as a distribution or core layer switch, the recommended software is Cisco IOS. Furthermore, at the time of writing, Catalyst 6500 IOS does not yet support AutoQoS nor the conditional-trust feature, therefore only CatOS examples of these models are given. However, both

AutoQoS VoIP and conditional trust have been committed for future releases of Catalyst 6500 IOS. The QoS design recommendations for a Catalyst 6500 switch at the access layer are summarized in Figure 2-22; the corresponding recommendations for Catalyst 6500s deployed in the distribution or core layers is shown in Figure 2-23.

Figure 2-22 Access Layer (CatOS) Catalyst 6500 QoS Design Options

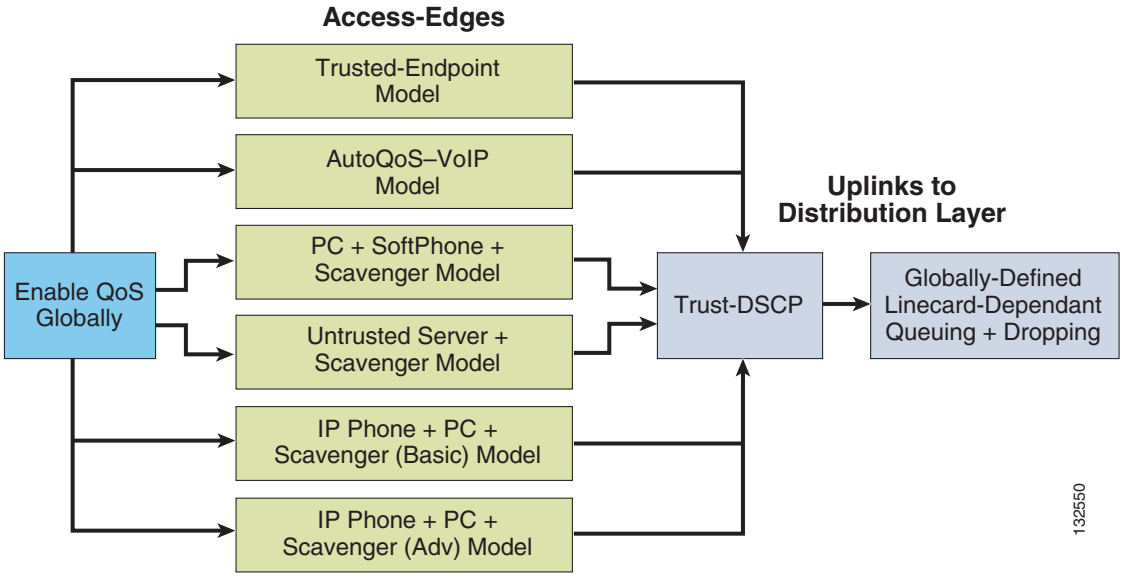
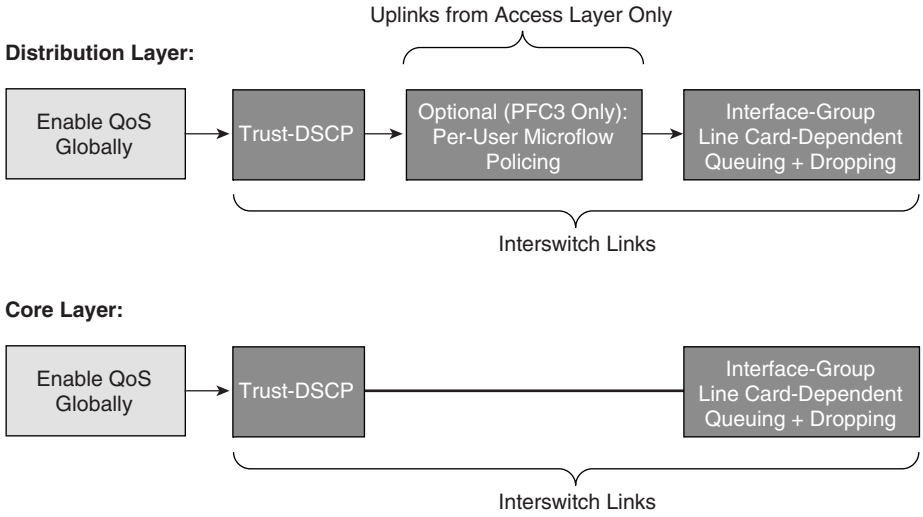


Figure 2-23 Distribution and/or Core Layer (IOS) Catalyst 6500 QoS Design



Note To narrow the scope of our discussion to the most current and relevant versions of the Catalyst 6500 switch-family, only the Catalyst 6500 with Supervisor 2 (PFC2) and Supervisor 720 (PFC3) be examined in this design chapter. For discussions of older versions of Catalyst 6000/6500s (such as Supervisor 1, 1a with/without a PFC), refer to the Cisco Press book, *Cisco Catalyst QoS: Quality of Service in Campus Networks*, by Mike Flannagan, Richard Froom and Kevin Turek.

QoS is globally disabled by default on Catalyst 6500s running either CatOS or IOS. When QoS is globally disabled, then all frames/packets that are passed-through the switch remain unaltered (which is equivalent to a trust CoS and trust DSCP state on all ports). When QoS is globally enabled, however, then all DSCP and CoS values are (by default) set to 0 (which is equivalent to an untrusted state on all ports).

The commands to enable and verify QoS globally on both CatOS and IOS Catalyst 6500s are shown below.

Example 2-46 Enabling QoS Globally on a Catalyst 6500—CatOS

```
CAT6500-PFC2-CATOS> (enable) set qos enable
QoS is enabled.
CAT6500-PFC2-CATOS> (enable)



CAT6500-PFC2-CATOS> (enable) show qos status
QoS is enabled on this switch.
CAT6500-PFC2-CATOS> (enable)
```

Example 2-47 Enabling QoS Globally on a Catalyst 6500—IOS

```
CAT6500-PFC2-IOS(config)#mls qos
CAT6500-PFC2-IOS(config)#end
CAT6500-PFC2-IOS#



CAT6500-PFC2-IOS#show mls qos
  QoS is enabled globally
  Microflow policing is enabled globally
Vlan or Portchannel(Multi-Earl) policies supported: Yes

 ----- Module [2] -----
  QoS global counters:
    Total packets: 65
    IP shortcut packets: 0
    Packets dropped by policing: 0
    IP packets with TOS changed by policing: 0
    IP packets with COS changed by policing: 0
    Non-IP packets with COS changed by policing: 0

CAT6500-PFC2-IOS#
```

Catalyst 6500—CatOS Defaults and Recommendations

CatOS specifies a number of default QoS settings per-port, which do not appear in the normal configuration output. However it is beneficial to be aware of what these defaults are and what they do, so as not to override them by mistake.

For example, CatOS allows the QoS policy-source to be defined by the local configuration or by the Common Open Policy Source (COPS) protocol, referring to a COPS Policy-Decision Point (PDP) Server. COPS is a QoS administration protocol that is both dynamic and scalable, but unfortunately it never gained mainstream acceptance. It is recommended to leave the switch's default policy-source as local (except of course in the extremely rare occurrence that COPS is actually deployed on the network).

Additionally, QoS policies may be applied to VLANs or to ports. There was never any significant advantage of using one base over the other; however, AutoQoS tools favor port-based QoS, as it is marginally simpler to configure. Port-based QoS is the default per-port setting and all examples in this chapter are configured using port-based QoS.

All ports (once QoS has been globally enabled) are set to an untrusted state by default. Also, by default, the trust-extension state is set to untrusted and the extended-CoS is correspondingly set to 0.

All packets received through an untrusted port (whether the untrusted port is the actual switch port or the extended switch port in the back of a Cisco IP Phone) are marked to a CoS value of 3, by default. This default marking should be set instead to 0 on all ports connected to untrusted endpoints by using the command **set port qos** *mod/port* **cos 0**. Furthermore, on all ports that are connected to conditionally-trusted endpoints (like Cisco IP Phones) it is recommended to use the command **set port qos** *mod/port* **cos 0** in conjunction with the command **set port qos** *mod/port* **cos-ext 0**.

It is recommended to leave all these port QoS settings at their defaults, with the exception of trust and cos/cos-ext—depending on the access edge model to be applied to the port, as is discussed in additional detail below.

Another Catalyst-OS default behavior to keep in mind is that ACLs and aggregate policers cannot be applied to more than one port in the same manner as these can when configured in IOS. For example, if an aggregate policer called **POLICE-VOIP** was defined to rate-limit flows to 128 kbps and if this policer were applied to two separate ports in CatOS, then it would rate limit flows from **both** ports to combined total of 128 kbps, instead of (the preferred behavior of) limiting flows to 128 kbps on a per-port basis (as is the case when configured in IOS). To work around this default behavior, ACLs and aggregate policers have to be uniquely defined on a per-port basis. To facilitate the administration of this additional configuration complexity, it is recommended that all CatOS ACLs and aggregate policers be defined with names that include the module and port they are to be applied to. For example, the previously defined aggregate policer **POLICE-VOIP** would become **POLICE-VOIP-3-1** when applied to port 3/1 and **POLICE-VOIP-3-2** when applied to port 3/2. This is the nomenclature adopted in the examples to follow in this chapter.

Note Administrators should keep in mind the maximum number of aggregate policers that can be configured via CatOS on a given Catalyst 6500 switch (currently 1023) when designing their access-edge policies. Depending on the chassis/linecard combination, this maximum number of aggregate policers may present scaling limitations to the advanced models presented in this design chapter.

Catalyst 6500—Trusted Endpoint Model

This section includes the following topics:

- Configuration
- Catalyst 6500 CatOS QoS Verification Commands

Configuration

For most Catalyst 6500 switch ports, setting the trust state to trust DSCP is a relatively straightforward command (in either CatOS or in IOS).

In this first example, DSCP trust is configured on a port in CatOS; in the second, DSCP trust is configured on a port/interface in IOS.

Example 2-48 Catalyst 6500 CatOS—Trusted Endpoint Model

```
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 trust trust-dscp
Port  3/1 qos set to trust-dscp.
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Commands

Catalyst 6500 CatOS QoS verification commands include the following:

- show port qos
- show qos status

show port qos

The Catalyst 6500 CatOS **show port qos** verification command returns the configured and runtime QoS states of a port. These may differ, as certain commands need to be committed (programmed into hardware) before they become effective.

In the example below, the switch has QoS globally enabled and the source of QoS policy decisions is the local configuration, as opposed to a Common-Open Policy Source Policy-Decision Point (COPS PDP).

Furthermore, the output shows that the port is configured for port-based QoS (by default) and has been set to trust DSCP from connected endpoints. No trust-extension has been configured, as this is not a conditionally-trusted endpoint model.

The output includes the linecard's queuing capabilities: 1P3Q1T (Transmit) and 1P1Q0T (Receive), as well as any ACLs that may be mapped to the port (however, no ACLs have been mapped to this port in this particular example).

Example 2-49 Show Port QoS Verification for a Catalyst 6500—CatOS Switch

```
CAT6500-PFC2-CATOS> (enable) show port qos 3/1
QoS is enabled for the switch.
QoS policy source for the switch set to local.

Port  Interface Type Interface Type Policy Source Policy Source
      config         runtime        config        runtime
----- -------------- -------------- ------------- -------------
 3/1      port-based     port-based          COPS         local

Port  TxPort Type  RxPort Type  Trust Type   Trust Type    Def CoS Def CoS
                                config       runtime       config  runtime
----- ------------ ------------ ------------ ------------- ------- -------
 3/1        1p3q1t       1p1q0t   trust-dscp   trust-dscp        0       0

Port  Ext-Trust Ext-Cos Trust-Device
----- --------- ------- ------------
 3/1  untrusted       0         none

(*)Runtime trust type set to untrusted.

Config:
Port  ACL name                          Type
----- --------------------------------- ----
No ACL is mapped to port 3/1.

Runtime:
Port  ACL name                          Type
----- --------------------------------- ----
```

```
No ACL is mapped to port 3/1.
CAT6500-PFC2-CATOS> (enable)
```

On non-GigabitEthernet linecards that use 2Q2T Transmit Queuing and 1Q4T Receive queuing (such as the WS-X6248-RJ-xx and WS-X6348-RJ-xx linecards), a hardware limitation prevents the proper functioning of port-based trust (which affects trust-cos, trust-ipprec, and trust-dscp). The **show port qos** command can be used to determine if the linecard is a 2Q2T-Tx/1Q4T-Rx linecard. These cards also listed in Table 2-5.

On such linecards, a workaround ACL can be used to achieve trust-functionality for trust-cos, trust-ipprec, and trust-dscp. The workaround ACL for trust-DSCP functionality on such linecards is shown below.

Example 2-50 Trust-DSCP Workaround ACL for Catalyst 6500 2Q2T-TX/1Q4T-Rx Non-Gigabit Linecards

```
CAT6500-PFC2-CATOS> (enable) set qos acl ip TRUST-DSCP trust-dscp any
TRUST-DSCP editbuffer modified. Use 'commit' command to apply changes.
CAT6500-PFC2-CATOS> (enable) commit qos acl TRUST-DSCP

QoS ACL 'TRUST-DSCP' successfully committed.
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl map TRUST-DSCP 4/1
```

Note To apply the QoS ACL you have defined (above), the ACL must be committed to hardware. The process of committing copies the ACL from a temporary editing buffer to the PFC hardware. Once resident in the PFC memory, the policy defined in the QoS ACL can be applied to all traffic that matches the Access Control Entries (ACEs). For ease of configuration, most administrators issue a **commit all** command, however you can commit a specific ACL (by name) to be sent from the editing buffer to PFC memory, as shown above.

In this second example, DSCP trust is configured on a port/interface in IOS.

Example 2-51 Catalyst 6500 IOS—Trusted Endpoint Example

```
CAT6500-PFC2-IOS(config)#interface FastEthernet3/1
CAT6500-PFC2-IOS(config-if)#mls qos trust dscp
```

Other Catalyst 6500 CatOS QoS verification commands include the following:

- show mls qos
- show mls qos interface

Note The ACL Trust workaround to the 2Q2T non-GigabitEthernet linecards (such as the such as the WS-X6248-RJ-xx and WS-X6348-RJ-xx) limitation of not supporting trust only applies in the access Layer of the campus (where CatOS is the recommended software for the Catalyst 6500). In the distribution and core layers, where IOS is the preferred software, all interfaces are recommended to be GigabitEthernet or higher.

Catalyst 6500 Auto QoS VoIP Model

At the time of writing, AutoQoS VoIP is only supported on the Catalyst 6500 in CatOS.

The AutoQoS VoIP macro for the Catalyst 6500 CatOS is divided into these two separate components:

- Global AutoQoS command (**set qos auto**)—Deals with all switch-wide related QoS settings that are not specific to any given interface, including CoS-to-queue maps, CoS-to-DSCP maps, and WRED settings for specific port types and global mappings.
- Port-specific automatic QoS command (**set port qos** *mod/port* **autoqos**)—Configures all inbound QoS parameters for a particular port to support IP Telephony devices.

The port-specific AutoQoS VoIP command for the Catalyst 6500 in CatOS supports the the following keyword options:

- **autoqos voip ciscoipphone**
- **autoqos voip ciscosoftphone**
- **autoqos trust [cos | dscp]**

The effects of enabling the Global AutoQoS command **set qos auto** on the Catalyst 6500 in CatOS are as follows.

Example 2-52 Catalyst 6500 Global AutoQoS Generated Configuration

```
set qos autoqos
---------------

set qos enable

set qos policy-source local
set qos ipprec-dscp-map 0 10 18 26 34 46 48 56
set qos cos-dscp-map 0 10 18 26 34 46 48 56
set qos dscp-cos-map 0-7:0 8-15:1 16-23:2 24-31:3 32-39:4 40-47:5 48-55:6 56-63:7
set qos acl default-action ip dscp 0
set qos map 2q2t tx queue 2 2 cos 5,6,7
set qos map 2q2t tx queue 2 1 cos 1,2,3,4
set qos map 2q2t tx queue 1 1 cos 0
set qos drop-threshold 2q2t tx queue 1 100 100
set qos drop-threshold 2q2t tx queue 2 80 100
set qos drop-threshold 1q4t rx queue 1 50 60 80 100
set qos txq-ratio 2q2t 80 20
set qos wrr 2q2t 100 255


set qos map 1p3q1t tx 1 1 cos 0
set qos map 1p3q1t tx 2 1 cos 1,2
set qos map 1p3q1t tx 3 1 cos 3,4
set qos map 1p3q1t tx 3 0 cos 6,7
set qos map 1p3q1t tx 4 cos 5
set qos wrr 1p3q1t 20 100 200
set qos wred 1p3q1t queue 1 70:100
set qos wred 1p3q1t queue 2 70:100
set qos wred 1p3q1t queue 3 70:90
set qos map 1p1q0t rx 1 cos 0,1,2,3,4
set qos map 1p1q0t rx 2 cos 5,6,7
set qos rxq-ratio 1p1q0t 80 20
set qos map 1p2q2t tx 1 2 cos 0
set qos map 1p2q2t tx 2 1 cos 1,2,3,4
set qos map 1p2q2t tx 2 2 cos 6,7
set qos map 1p2q2t tx 3 cos 5
set qos txq-ratio 1p2q2t 75 15 15
set qos wrr 1p2q2t 50 255
set qos wred 1p2q2t queue 1 1 40:70
set qos wred 1p2q2t queue 1 2 70:100
set qos wred 1p2q2t queue 2 1 40:70
```

```
set qos wred 1p2q2t queue 2 2 70:100
set qos map 1p1q4t rx 1 1 cos 0
set qos map 1p1q4t rx 1 3 cos 1,2,3,4
set qos map 1p1q4t rx 1 4 cos 6,7
set qos map 1p1q4t rx 2 cos 5
set qos drop-threshold 1p1q4t rx queue 1 50 60 80 100

set qos map 1p2q1t tx 1 1 cos 0
set qos map 1p2q1t tx 2 1 cos 1,2,3,4
set qos map 1p2q1t tx 2 cos 6,7
set qos map 1p2q1t tx 3 cos 5
set qos txq-ratio 1p2q1t 75 15 15
set qos wrr 1p2q1t 50 255
set qos wred 1p2q1t queue 1 70:100
set qos wred 1p2q1t queue 2 70:100
set qos map 1p1q8t rx 1 1 cos 0
set qos map 1p1q8t rx 1 5 cos 1,2
set qos map 1p1q8t rx 1 8 cos 3,4
set qos map 1p1q8t rx 2 cos 5,6,7
set qos wred 1p1q8t queue 1 1 40:70
set qos wred 1p1q8t queue 1 5 60:90
set qos wred 1p1q8t queue 1 8 70:100
set qos rxq-ratio 1p1q8t 80 20

set qos policed-dscp-map 0:0
set qos policed-dscp-map 1:1
set qos policed-dscp-map 2:2
```

<repetitive output truncated>

```
set qos policed-dscp-map 61:61
set qos policed-dscp-map 62:62
set qos policed-dscp-map 63:63


set qos policed-dscp-map excess-rate 0:0
set qos policed-dscp-map excess-rate 1:1
set qos policed-dscp-map excess-rate 2:2
```

<repetitive output truncated>

```
set qos policed-dscp-map excess-rate 61:61
set qos policed-dscp-map excess-rate 62:62
set qos policed-dscp-map excess-rate 63:63
```

The effects of enabling the port-specific **set port qos** *mod/port* **autoqos voip ciscoipphone** command on the Catalyst 6500 in CatOS are as follows.

Example 2-53 Catalyst 6500 Port-Specific AutoQoS VoIP CiscoIPPhone Generated Configuration

```
set port qos mod/port autoqos voip ciscoipphone
---------------
set port qos mod/port policy-source local
set port qos mod/port port-based
set port qos mod/port cos 0
set port qos mod/port cos-ext 0
set port qos mod/port trust-ext untrusted
set port qos mod/port trust-device ciscoipphone
```

If the port is on a 2Q2T-Tx/1Q4T-Rx (non-GigabitEthernet) linecard, the configuration is as follows:

```
set qos acl ip ACL_IP-PHONES trust-cos any
```

```
commit qos acl ACL_IP-PHONES
set qos acl map ACL_IP-PHONES mode/port
set port qos mod/port trust trust-cos
```

If the port type is another port type, the configuration is as follows:

```
set port qos mod/port trust trust-cos
```

The effects of enabling the port-specific set port qos mod/port autoqos voip ciscosoftphone command on the Catalyst 6500 in CatOS are as follows.

Example 2-54 Catalyst 6500 Port-Specific AutoQoS VoIP CiscoSoftPhone Generated Configuration

```
set port qos mod/port autoqos voip ciscosoftphone
---------------
set port qos mod/port policy-source local
set port qos mod/port port-based
set port qos mod/port cos 0
set port qos mod/port cos-ext 0
set port qos mod/port trust-ext untrusted
set port qos mod/port trust-device none
set port qos mod/port trust untrusted
set qos policer aggregate POLICE_SOFTPHONE-DSCP46-mod-port rate 320 burst 20 policed-dscp
set qos policer aggregate POLICE_SOFTPHONE-DSCP26-mod-port rate 32  burst 8 policed-dscp
set qos acl ip ACL_IP-SOFTPHONE-mod-port trust-dscp aggregate
POLICE_SOFTPHONE-DSCP46-mod-port any dscp-field 46
set qos acl ip ACL_IP-SOFTPHONE-mod-port trust-dscp aggregate
POLICE_SOFTPHONE-DSCP26-mod-port any dscp-field 26
commit qos acl ACL_IP-SOFTPHONE-mod-port
set qos acl map ACL_IP-SOFTPHONE-mod-port mod/port
```

The effects of enabling the port-specific **set port qos** *mod/port* **autoqos voip trust cos** command on the Catalyst 6500 in CatOS are as follows.

Example 2-55 Catalyst 6500 Port-Specific AutoQoS VoIP Trust CoS Generated Configuration

```
set port qos mod/port autoqos trust cos
---------------
set port qos mod/port policy-source local
set port qos mod/port port-based
set port qos mod/port cos 0
set port qos mod/port cos-ext 0
set port qos mod/port trust-ext untrusted
set port qos mod/port trust-device none
```

If the port is on a 2Q2T-Tx/1Q4T-Rx (non-GigabitEthernet) linecard, the configuration is as follows:

```
set qos acl ip ACL_IP-TRUSTCOS trust-cos any
commit qos acl ACL_IP-TRUSTCOS
set qos acl map ACL_IP-TRUSTCOS mode/port
set port qos mod/port trust trust-cos
```

If the port type is another port type, the configuration is as follows:

```
set port qos mod/port trust trust-cos
```

The AutoQoS VoIP command with the Trust-DSCP option is virtually identical to the Trust-CoS option (albeit the final configuration commands are **set port qos** *mod/port* **trust trust-dscp** instead of **trust-cos**).

Catalyst 6500—Untrusted PC + SoftPhone with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst 6500 CatOS QoS Verification Commands

Configuration

The radical difference in syntax between CatOS and IOS becomes increasingly apparent as more complex access edge models are presented.

In the Untrusted PC + SoftPhone + Scavenger Model, shown below, per-application aggregate policers are defined—one each for SoftPhone VoIP traffic, SoftPhone call signaling traffic, and PC Data traffic. Then an ACL (titled "SOFTPHONE-PC-mod-port") with multiple ACEs is defined, with each ACE referencing its associated aggregate policer. Once complete, the ACL is committed to PFC memory and then mapped to the desired switch port(s). Switch responses to the commands have been omitted to simplify the example.

Example 2-56 Catalyst 6500 CatOS—Untrusted PC + SoftPhone + Scavenger Model Configuration

```
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map 0,24,46:8
! Excess traffic marked DSCP 0 or CS3 or EF will be remarked to CS1
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate SOFTPHONE-VOICE-3-1
 rate 128 burst 8000 policed-dscp
! Defines the policer for SoftPhone VoIP traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate SOFTPHONE-SIGNALING-3-1
 rate 32 burst 8000 policed-dscp
! Defines the policer for SoftPhone call signaling traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate PC-DATA-3-1
 rate 5000 burst 8000 policed-dscp
! Defines the policer for PC Data traffic
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip SOFTPHONE-PC-3-1 dscp 46
 aggregate SOFTPHONE-VOICE-3-1 udp any any range 16384 32767
! Binds ACL to policer and marks in-profile SoftPhone VoIP to DSCP EF
CAT6500-PFC2-CATOS> (enable) set qos acl ip SOFTPHONE-PC-3-1 dscp 24
 aggregate SOFTPHONE-SIGNALING-3-1 tcp any any range 2000 2002
! Binds ACL to policer marks in-profile call signaling to DSCP CS3
CAT6500-PFC2-CATOS> (enable) set qos acl ip SOFTPHONE-PC-3-1 dscp 0
 aggregate PC-DATA-3-1 any
! Binds ACL to policer and marks in-profile PC Data traffic to DSCP 0
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) commit qos acl SOFTPHONE-PC-3-1
    ! Commits ACL to PFC hardware
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 cos 0
    ! Sets CoS to 0 for all untrusted packets
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 trust untrusted
    ! Sets the port trust state to untrusted
CAT6500-PFC2-CATOS> (enable) set qos acl map SOFTPHONE-PC-3-1 3/1
    ! Attaches ACL to switch port
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS verification commands for the Untrusted PC + Softphone + Scavenger model include the following:

- show qos status

- show qos maps
- show port qos
- show qos acl
- show qos policer
- show qos statistics

show qos maps

The Catalyst 6500 CatOS **show qos maps** verification command is fairly similar to the **show mls qos maps** MLS QoS verification command. It returns the configured CoS-DSCP, IPPrec-DSCP, DSCP-CoS, and Normal-Rate and Excess-Rate Policed-DSCP Maps. The command can return configured maps (which may or may not be committed to the PFC) or runtime maps.

In the (truncated) runtime example below, all maps are at their default states, with the exception of the Normal-Rate Policed-DSCP map, which has DSCP 0, CS3 (24), and EF (46) mapped for out-of-profile markdown to DSCP CS1 (8).

Example 2-57 Show QoS Maps Verification for Catalyst 6500 Switch—CatOS

```
CAT6500-PFC2-CATOS> (enable) show qos maps runtime
CoS - DSCP map:
CoS   DSCP
---   ----
  0   0
  1   8
  2   16
  3   24
  4   32
  5   40
  6   48
  7   56

IP-Precedence - DSCP map:
IP-Prec   DSCP
-------   ----
      0   0
      1   8
      2   16
      3   24
      4   32
      5   40
      6   48
      7   56

DSCP - CoS map:
DSCP                              CoS
--------------------------------  ---
                               0-7  0
                              8-15  1
                             16-23  2
                             24-31  3
                             32-39  4
                             40-47  5
                             48-55  6
                             56-63  7

DSCP - Policed DSCP map normal-rate:
DSCP                              Policed DSCP
--------------------------------  ------------
```

```
                                1  1
                                2  2
                                3  3
                                4  4
                                5  5
                                6  6
                                7  7
                          0,8,24,46  8
                                9  9
                               10  10
<output truncated>
                               63  63
DSCP - Policed DSCP map excess-rate:
DSCP                               Policed DSCP
-------------------------------  ------------
                                0  0
                                1  1
                                2  2
                                3  3
                                4  4
                                5  5
<output truncated>
                               63  63

CAT6500-PFC2-CATOS> (enable)
```

show qos acl

The Catalyst 6500 CatOS **show qos acl** verification command returns information about ACLs and ACEs that have been configured for QoS purposes. ACL information can be displayed for configuration ACLs or runtime ACLs.

In the example below, three variations of the **show qos acl** command are displayed.

The first one displays the QoS ACLs that are still in the edit-buffer and indicates whether or not the ACL(s) have been committed to PFC hardware. In this example, the ACL "SOFTPHONE-PC-3-1" has been committed to the PFC.

The second example displays runtime ACE level information for a given ACL (or all ACLs, if the keyword **all** is used instead of the ACL name). Each ACE's DSCP markings, associated aggregate policer, and filtering criteria are displayed.

The third example displays the VLANs and/or ports that the ACL has been applied to. In this specific example, port 3/1 has the SOFTPHONE-PC-3-1 ACL applied to it.

Example 2-58 Show QoS ACL Verification for Catalyst 6500 Switch—CatOS

```
CAT6500-PFC2-CATOS> (enable) show qos acl editbuffer
ACL                              Type Status
-------------------------------  ---- ----------
SOFTPHONE-PC-3-1 IP   Committed
CAT6500-PFC2-CATOS> (enable)

CAT6500-PFC2-CATOS> (enable) show qos acl info runtime SOFTPHONE-PC-3-1

set qos acl IP SOFTPHONE-PC-3-1
---------------------------------------------
1. dscp 46 aggregate SOFTPHONE-VOICE-PC-3-1 udp any any range 16384 32767
2. dscp 24 aggregate SOFTPHONE-SIGNALING-PC-3-1 tcp any any range 2000 2002
3. dscp 0 aggregate PC-DATA-PC-3-1 any
CAT6500-PFC2-CATOS> (enable)
```

```
CAT6500-PFC2-CATOS> (enable) show qos acl map runtime SOFTPHONE-PC-3-1
QoS ACL mappings on input side:
ACL name                          Type Vlans
-------------------------------- ---- --------------------------------
SOFTPHONE-PC-3-1               IP
ACL name                          Type Ports
-------------------------------- ---- --------------------------------
SOFTPHONE-PC-3-1               IP 3/1
CAT6500-PFC2-CATOS> (enable)
```

show qos policer

The Catalyst 6500 CatOS **show qos policer** verification command displays the Normal and Excess Rates and Burst for any/all policers.

In the example below, three aggregate policers have been defined: SOFTPHONE-VOICE-3-1, SOFTPHONE-SIGNALING-3-1, and PC-DATA-3-1, with Normal Rates of 128 kbps, 32 kbps, and 5 Mbps, respectively. Each of these policers will markdown excess traffic according to the Policed-DSCP Normal-Rate and are attached to the ACL: SOFTPHONE-PC-3-1.

Example 2-59 Show QoS Policer Verification for Catalyst 6500 Switch—CatOS

```
CAT6500-PFC2-CATOS> (enable) show qos policer runtime all
Warning: Runtime information may differ from user configured setting due to hard
ware granularity.
QoS microflow policers:
QoS aggregate policers:
Aggregate name        Avg. rate (kbps) Burst size (kb) Normal action
--------------------- ---------------- --------------- -------------
SOFTPHONE-VOICE-3-1           128            7936  policed-dscp
                      Excess rate (kbps) Excess burst size (kb) Excess action
                      ------------------ ---------------------- -------------
                                31457280                  31744  policed-dscp
                                ACL attached
                                -----------------------------------
                                SOFTPHONE-PC-3-1

Aggregate name        Avg. rate (kbps) Burst size (kb) Normal action
--------------------- ---------------- --------------- -------------
SOFTPHONE-SIGNALING-3-1        32            7936 policed-dscp
                      Excess rate (kbps) Excess burst size (kb) Excess action
                      ------------------ ---------------------- -------------
                                31457280                  31744  policed-dscp
                                ACL attached
                                -----------------------------------
                                SOFTPHONE-PC-3-1

Aggregate name        Avg. rate (kbps) Burst size (kb) Normal action
--------------------- ---------------- --------------- -------------
PC-DATA-3-1                   4864            7936  policed-dscp
                      Excess rate (kbps) Excess burst size (kb) Excess action
                      ------------------ ---------------------- -------------
                                31457280                  31744  policed-dscp
                                ACL attached
                                -----------------------------------
                                SOFTPHONE-PC-3-1

CAT6500-PFC2-CATOS> (enable)
```

show qos statistics

The Catalyst 6500 CatOS **show qos statistics** verification command displays various dynamic statistics regarding the QoS policies.

In the three part example below, the first variation of the command **show qos statistics <mod/port>** returns queuing statistics for the port. Specifically it reports any drops due to queue buffer overfill and breaks these drops down by queues/thresholds—depending on the queuing structure of the module. In this first part of the example, no drops have occurred due to queuing buffer overfills.

In the second example, aggregate policing statistics are displayed via the **show qos statistics aggregate-policer** variation of the command. The number of packets conforming or exceeding a given policer is reported. In this part of the example, the command reports that no packets have exceeded the SOFTPHONE-VOIP-3-1 or SOFTPHONE-SIGNALING-3-1 policer, but a few have exceeded the PC-DATA-3-1 policer.

And finally, in the third example, the number of packets that have been dropped due to policing and/or remarked at Layer 3 or Layer 2 are reported with the **show qos statistics l3stats** command. In this final part of the example, the command reports that no packets have been dropped due to policing, but thousands of packets have had their Layer 3 and Layer 2 markings modified by the configured policy.

Example 2-60 Show QoS Statistics Verification for Catalyst 6500 Switch—CatOS

```
CAT6500-PFC2-CATOS> (enable) show qos statistics 3/1
Tx port type of port 3/1 : 1p3q1t
WRED and tail drops are accumulated in one counter per queue.
Q #  Packets dropped
---  ----------------------------------------------
1    0 pkts
2    0 pkts
3    0 pkts
4    0 pkts

Rx port type of port 3/1 : 1p1q0t
For untrusted ports all the packets are sent to the same queue,
Rx thresholds are disabled, tail drops are reported instead.
Q #  Threshold #:Packets dropped
---  ----------------------------------------------
1    0:0 pkts
2    0:0 pkts

CAT6500-PFC2-CATOS> (enable)

CAT6500-PFC2-CATOS> (enable) show qos statistics aggregate-policer
QoS aggregate-policer statistics:
Aggregate policer               Allowed packet Packets exceed
                                count          excess rate
------------------------------- -------------- --------------
SOFTPHONE-VOICE-3-1                    27536              0
SOFTPHONE-SIGNALING-3-1                224              0
PC-DATA-3-1                            470069            645
CAT6500-PFC2-CATOS> (enable)

CAT6500-PFC2-CATOS> (enable) show qos statistics l3stats
Packets dropped due to policing:                0
IP packets with ToS changed:               169286
IP packets with CoS changed:                83507
Non-IP packets with CoS changed:                0
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500—Untrusted Server with Scavenger-Class QoS Model

This section includes the following topics:

- Configuration
- Catalyst 6500 CatOS QoS Verification Commands

Additional flexibility is offered to the Untrusted Server + Scavenger Model due to the Catalyst 6500 PFC2/PFC3's support of Dual-Rate Policing (as described in RFC 2698 "A Two Rate Three Color Marker" and as illustrated in Figure 4-5).

Using a Dual-Rate policer, three colors are used to indicate:

- *Conforming traffic* (within the normal rate)
- *Excess traffic* (exceeding the normal rate but less than the excess rate)
- *Violating traffic* (exceeding both the normal and excess rate)

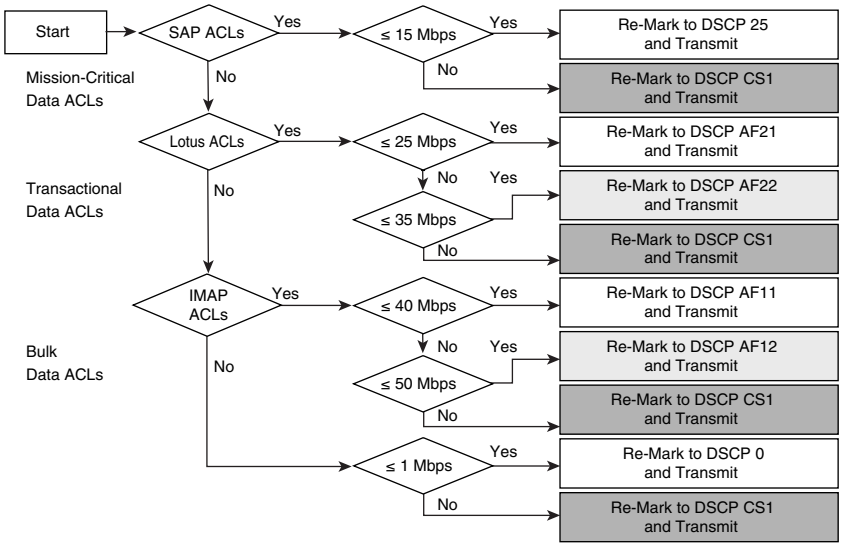
The Dual-Rate policer is intended to complement the RFC 2597 Assured Forwarding Groups Diff-Serv marking scheme. To illustrate, consider Transactional Data traffic, which is marked to AF Class 2. Conforming Transactional Data should be marked to AF21, Excess Transactional Data traffic should be marked-down to AF22, and Violating Transactional Data traffic should be marked-down further to AF23.

Such a markdown scheme is intended to be complemented further by DSCP-based WRED congestion avoidance. In this manner, in the event of congestion, AF23 is dropped more aggressively than AF22 which, in turn, is dropped more aggressively than AF21.

However, since Catalyst 6500 queuing and congestion-avoidance is determined primary by CoS markings, the standards-based DSCP model cannot be followed completely at this time on this platform (since AF21/AF22/AF23 all share CoS 3, this does not allow for granular sub-class QoS). Therefore, the use of Scavenger class markings for violating traffic could be used to achieve a similar overall effect, while maintaining consistency with QoS designs previously presented for other Catalyst platforms.

Under such a modified Untrusted Multi-Application Server with Scavenger-Class QoS Model, Excess Transactional Data traffic can be marked down to AF22 and Violating Transactional Data traffic can be marked down to DSCP CS1 (Scavenger). Similarly, Excess Bulk Data traffic can be marked down to AF12 and Violating Bulk Data traffic can be marked down to DSCP CS1 (Scavenger). This modified model is shown in Figure 2-24.

Figure 2-24 Catalyst 6500 PFC2/PFC3 Untrusted Endpoint—Multi-Application Server with Scavenger-Class QoS (Dual-Rate Policing) Model



Configuration

The configuration for a Catalyst 6500 CatOS Untrusted Endpoint Dual-Rate Policing of a Multi-Application Server with Scavenger-Class QoS example is shown below.

Example 2-61 Catalyst 6500 CatOS—Untrusted Multi-Application Server with Scavenger-Class QoS (Dual-Rate Policing) Model Configuration

```
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map normal-rate 0,25:8
    ! Excess SAP and Data traffic is marked down to DSCP CS1 (Scavenger)
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map normal-rate 18:20
    ! Excess Transactional Data traffic is marked down from DSCP AF21 to AF22
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map excess-rate 18:8
    ! Violating Transactional Data traffic is marked down to CS1
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map normal-rate 10:12
    ! Excess Bulk Data traffic is marked down from DSCP AF11 to AF12
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map excess-rate 10:8
! Violating Bulk Data traffic is marked down to CS1
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate SAP-3-1
rate 15000 burst 8000 policed-dscp
! Defines the policer for Mission-Critical Data (SAP) traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate LOTUS-3-1
rate 25000 policed-dscp erate 35000 policed-dscp burst 8000
! Defines the dual-rate policer for Transactional Data (Lotus) traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate IMAP-3-1
rate 40000 policed-dscp erate 50000 policed-dscp burst 8000
! Defines the dual-rate policer for Bulk Data (IMAP) traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate DATA-3-1
rate 1000 burst 8000 policed-dscp
! Defines the policer for other data traffic
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip UNTRUSTED-SERVER-3-1 dscp 25
aggregate SAP-3-1 tcp any range 3200 3203 any
! Binds ACL to policer and marks in-profile SAP to DSCP 25
CAT6500-PFC2-CATOS> (enable) set qos acl ip UNTRUSTED-SERVER-3-1 dscp 25
```

```
aggregate SAP-3-1 tcp any eq 3600 any
! Binds ACL to policer and marks in-profile SAP to DSCP 25
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip UNTRUSTED-SERVER-3-1 dscp 18
aggregate LOTUS-3-1 tcp any eq 1352 any
! Binds ACL to dual-rate policer and marks in-profile Lotus to DSCP AF21
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip UNTRUSTED-SERVER-3-1 dscp 10
aggregate IMAP-3-1 tcp any eq 143 any
! Binds ACL to dual-rate policer and marks in-profile IMAP to DSCP AF11
CAT6500-PFC2-CATOS> (enable) set qos acl ip UNTRUSTED-SERVER-3-1 dscp 10
aggregate IMAP-3-1 tcp any eq 220 any
! Binds ACL to dual-rate policer and marks in-profile IMAP to DSCP AF11
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip UNTRUSTED-SERVER-3-1 dscp 0
aggregate DATA-3-1 any
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) commit qos acl UNTRUSTED-SERVER-3-1
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 cos 0
    ! Sets CoS to 0 for all untrusted packets
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 trust untrusted
    ! Sets the port trust state to untrusted
CAT6500-PFC2-CATOS> (enable) set qos acl map UNTRUSTED-SERVER-3-1 3/1
    ! Attaches ACL to switch port
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Commands

Catalyst 6500 CatOS QoS verification commands for the Untrusted Server + Scavenger model include the following:

- show qos status
- show qos maps
- show port qos
- show qos acl
- show qos policer
- show qos statistics

Catalyst 6500—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Basic) Model

This section includes the following topics:

- Configuration
- Catalyst 6500 CatOS QoS Verification Commands

Configuration

In the Conditionally-Trusted IP Phone + PC + Scavenger (Basic) Model for the Catalyst 6500 (CatOS), four aggregate policers are defined, one each for Voice from the VVLAN, call signaling from the VVLAN, all other traffic from the VVLAN, and for all PC Data traffic. Conditional trust is extended to the IP Phones via the trust-device command, as shown below.

Example 2-62 Catalyst 6500 CatOS—Conditionally-Trusted IP Phone + PC + Scavenger (Basic) Model Configuration

```
CAT6500-PFC2-CATOS> (enable) set qos cos-dscp-map 0 8 16 24 32 46 48 56
    ! Modifies default CoS-DSCP mapping so that CoS 5 is mapped to DSCP EF
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map 0,24:8
! Excess traffic marked DSCP 0 or CS3 is remarked to CS1
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate VVLAN-VOICE-3-1
rate 128 burst 8000 drop
! Defines the policer for IP Phone VoIP traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate VVLAN-SIGNALING-3-1
rate 32 burst 8000 policed-dscp
! Defines the policer for IP Phone call signaling traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate VVLAN-ANY-3-1
rate 32 burst 8000 policed-dscp
! Defines the policer for any other traffic sourced from the VVLAN
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate PC-DATA-3-1
rate 5000 burst 8000 policed-dscp
! Defines the policer for PC Data traffic
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-BASIC-3-1 dscp 46
aggregate VVLAN-VOICE-3-1 udp 10.1.110.0 0.0.0.255 any range 16384 32767
! Binds ACL to policer and marks in-profile VVLAN VoIP to DSCP EF
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-BASIC-3-1 dscp 24
aggregate VVLAN-SIGNALING-3-1 udp 10.1.110.0 0.0.0.255 any range 2000 2002
! Binds ACL to policer marks in-profile VVLAN call signaling to DSCP CS3
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-BASIC-3-1 dscp 0
aggregate VVLAN-ANY-3-1 10.1.110.0 0.0.0.255
! Binds ACL to policer and marks all other VVLAN traffic to DSCP 0
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-BASIC-3-1 dscp 0
aggregate PC-DATA-3-1 any
! Binds ACL to policer and marks in-profile PC Data traffic to DSCP 0
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) commit qos acl IPPHONE-PC-BASIC-3-1
    ! Commits ACL to PFC hardware
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 cos 0
    ! Sets CoS to 0 for all untrusted packets (when there is no IP Phone on the port)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 cos-ext 0
    ! Sets CoS to 0 for all untrusted PC-generated packets (behind an IP Phone)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 trust-ext untrusted
    ! Ignore any CoS values for all PC-generated packets (behind an IP Phone)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 trust-device ciscoipphone
    ! Conditional trust (for Cisco IP Phones only)
CAT6500-PFC2-CATOS> (enable) set qos acl map IPPHONE-PC-BASIC-3-1 3/1
    ! Attaches ACL to switch port
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Commands

Catalyst 6500 CatOS QoS verification commands for the Conditionally-Trusted IP Phone + PC + Scavenger (Basic) Model

include the following:

- show qos status
- show qos maps
- show port qos
- show qos acl
- show qos policer
- show qos statistics

Note As previously mentioned, on non-GigabitEthernet linecards that use 2Q2T Transmit Queuing and 1Q4T Receive queuing (such as the WS-X6248-RJ-xx and the WS-X6348-RJ-xx), a hardware limitation prevents the proper functioning of port-based trust (which affects trust-cos, trust-ipprec, and trust-dscp). On such linecards, a workaround ACL can be used to achieve trust functionality. For such an example, see the Catalyst 6500 Trusted Endpoint Model section (Example 2-50) of this chapter.

Catalyst 6500—Conditionally-Trusted IP Phone + PC with Scavenger-Class QoS (Advanced) Model

This section includes the following topics:

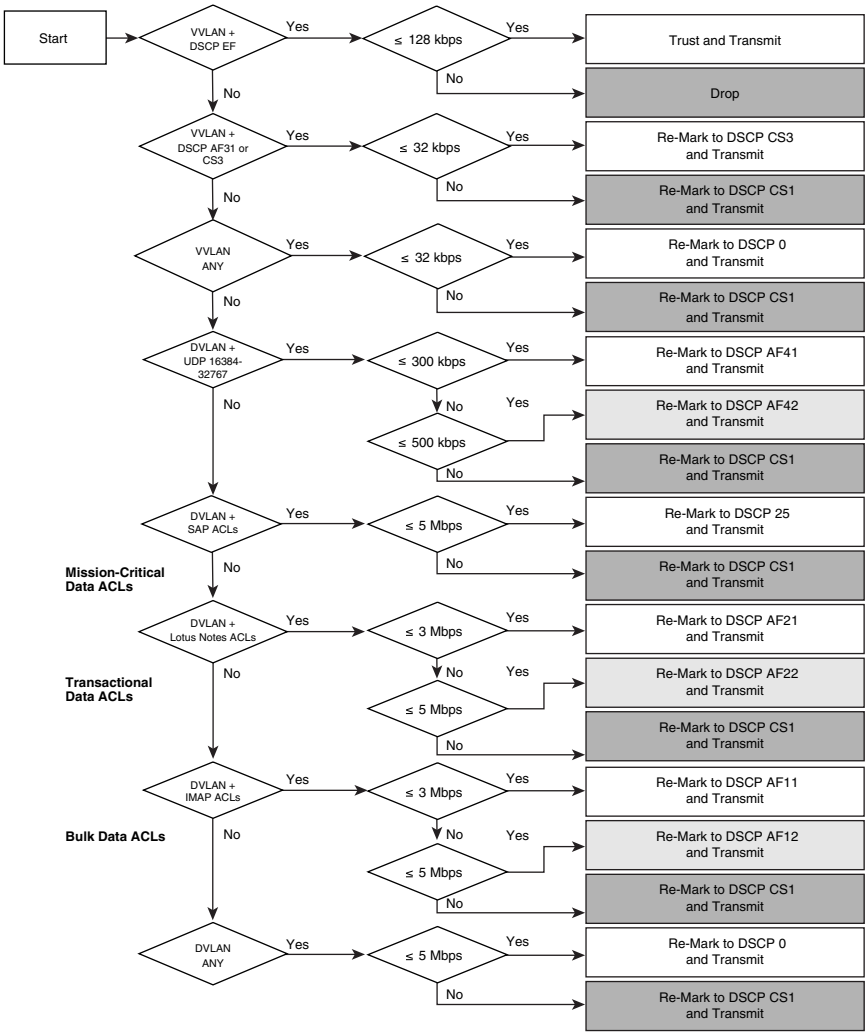
- Configuration
- Catalyst 6500 CatOS QoS Verification Commands

The Catalyst 6500 Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model leverages the Dual-Rate Policing capabilities of the PFC2/PFC3. This feature is described in the Catalyst 6500 Untrusted Server + Scavenger section of this chapter (which applies the feature to Server-to-Client flows). In this example, the Dual-Rate Policing feature is applied to Client-to-Server flows to complement the Untrusted Server + Scavenger Model.

Dual-Rate Policing, in this context, allows for graduated markdown of Interactive-Video (from PCs), Transactional Data, and Bulk Data. Specifically, in this example Interactive-Video is marked down to AF42 if it is in excess of 300 kbps but less than 500 kbps; if it is greater than 500 kbps, it is marked down to Scavenger (CS1). Similarly, Transactional Data and Bulk Data are marked down to AF22 and AF12 (respectively) if they are in excess of 3 Mbps but less than 5 Mbps; if they are in excess of 5 Mbps, then they are both marked down to Scavenger (CS1). All other policers are consistent with the single-rate policer model.

The Catalyst 6500 PFC2/PFC3 Conditionally-Trusted Endpoint Dual-Rate Policing—IP Phone + PC + Scavenger (Advanced) Model is illustrated in Figure 2-25.

Figure 2-25 Catalyst 6500 PFC2/PFC3 Conditionally-Trusted Endpoint Dual-Rate Policing—IP Phone + PC + Scavenger (Advanced) Model



Legend for Figure 2-25:

- MCD = Mission Critical Data
- TD = Transactional Data
- BD = Bulk Data

Note The discrete traffic watermarks at which graduated markdown should occur are at the network administrator's discretion and will vary from enterprise to enterprise and application to application.

Configuration

An example configuration for a Catalyst 6500 CatOS Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model is shown below.

Example 2-63 Catalyst 6500 CatOS—Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) Model Configuration

```
CAT6500-PFC2-CATOS> (enable) set qos cos-dscp-map 0 8 16 24 32 56 48 56
    ! Modifies default CoS-DSCP mapping so that CoS 5 is mapped to DSCP EF
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map normal-rate 0,24,25:8
    ! Excess Data, call signaling and MC-Data traffic is marked down to CS1
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map normal-rate 10:12
    ! Excess Bulk traffic is marked down from DSCP AF11 to AF12
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map excess-rate 10:8
    ! Violating Bulk traffic is marked down to DSCP CS1
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map normal-rate 18:20
    ! Excess Transactional Data traffic is marked down from AF21 to AF22
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map excess-rate 18:8
    ! Violating Transactional Data traffic is marked down to DSCP CS1
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map normal-rate 34:36
    ! Excess Interactive-Video traffic is marked down from AF41 to AF42
CAT6500-PFC2-CATOS> (enable) set qos policed-dscp-map excess-rate 34:8
    ! Violating Interactive-Video traffic is marked down to DSCP CS1
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate VVLAN-VOICE-3-1
rate 128 burst 8000 drop
! Defines the policer for IP Phone VoIP traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate VVLAN-SIGNALING-3-1
rate 32 burst 8000 policed-dscp
! Defines the policer for IP Phone call signaling traffic
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate VVLAN-ANY-3-1
rate 32 burst 8000 policed-dscp
! Defines the policer for any other traffic sourced from the VVLAN
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate PC-VIDEO-3-1
rate 300  policed-dscp erate 500 policed-dscp burst 8000
! Defines the Dual-Rate policer for Interactive-Video
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate MISSION-CRITICAL-3-1
rate 5000 burst 8000 policed-dscp
! Defines the policer for Mission-Critical Data
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate TRANSACTIONAL-3-1
rate 3000 policed-dscp erate 5000 policed-dscp burst 8000
! Defines the Dual-Rate policer for Transactional Data
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate BULK-3-1
rate 3000 policed-dscp erate 5000 policed-dscp burst 8000
! Defines the Dual-Rate policer for Bulk Data
CAT6500-PFC2-CATOS> (enable) set qos policer aggregate PC-DATA-3-1
rate 5000 burst 8000 policed-dscp
! Defines the policer for all other PC Data traffic
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 46
aggregate VVLAN-VOICE-3-1 udp 10.1.110.0 0.0.0.255 any range 16384 32767
! Binds ACL to policer and marks in-profile VVLAN VoIP to DSCP EF
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 24
aggregate VVLAN-SIGNALING-3-1 tcp 10.1.110.0 0.0.0.255 any range 2000 2002
! Binds ACL to policer marks in-profile VVLAN call signaling to DSCP CS3
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 0
aggregate VVLAN-ANY-3-1 10.1.110.0 0.0.0.255
! Binds ACL to policer and marks all other VVLAN traffic to DSCP 0
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 34
aggregate PC-VIDEO-3-1 udp any any range 16384 32767
! Binds ACL to Dual-Rate policer and marks in-profile PC Video to AF41
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 25
aggregate MISSION-CRITICAL-3-1 tcp any any range 3200 3203
! Binds ACL to policer and marks in-profile SAP to DSCP 25
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 25
aggregate MISSION-CRITICAL-3-1 tcp any any eq 3600
```

```
! Binds ACL to policer and marks in-profile SAP to DSCP 25
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 18
aggregate TRANSACTIONAL-3-1 tcp any any eq 1352
! Binds ACL to Dual-Rate policer and marks in-profile Lotus to AF21
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 10
aggregate BULK-3-1 tcp any any eq 143
! Binds ACL to Dual-Rate policer and marks in-profile IMAP to AF11
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 10
aggregate BULK-3-1 tcp any any eq 220
! Binds ACL to Dual-Rate policer and marks in-profile IMAP to AF11
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos acl ip IPPHONE-PC-ADVANCED-3-1 dscp 0
aggregate PC-DATA-3-1 any
! Binds ACL to policer and marks other in-profile PC data to DSCP 0
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) commit qos acl IPPHONE-PC-ADVANCED-3-1
    ! Commits ACL to PFC hardware
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 cos 0
    ! Sets CoS to 0 for all untrusted packets (when there is no IP Phone on the port)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 cos-ext 0
    ! Sets CoS to 0 for all untrusted PC-generated packets (behind an IP Phone)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 trust-ext untrusted
    ! Ignore any CoS values for all PC-generated packets (behind an IP Phone)
CAT6500-PFC2-CATOS> (enable) set port qos 3/1 trust-device ciscoipphone
    ! Conditional trust (for Cisco IP Phones only)
CAT6500-PFC2-CATOS> (enable) set qos acl map IPPHONE-PC-ADVANCED-3-1 3/1
    ! Attaches ACL to switch port
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Commands

Catalyst 6500 CatOS QoS verification commands for the Conditionally-Trusted IP Phone + PC + Scavenger (Advanced) model include the following:

- show qos status
- show qos maps
- show port qos
- show qos acl
- show qos policer
- show qos statistics

Note As previously mentioned, on non-GigabitEthernet linecards that use 2Q2T Transmit Queuing and 1Q4T Receive queuing (such as the WS-X6248-RJ-xx and the WS-X6348-RJ-xx), a hardware limitation prevents the proper functioning of port-based trust (which affects trust-cos, trust-ipprec, and trust-dscp). On such linecards, a workaround ACL can be used to achieve trust functionality. For such an example, see Catalyst 6500—Trusted Endpoint Model (Example 2-50).

Catalyst 6500—Queuing and Dropping

This section includes the following topics:

- Catalyst 6500 Queuing and Dropping Overview
- Catalyst 6500 Transmit Queuing and Dropping Linecard Options
- Catalyst 6500—2Q2T Queuing and Dropping
- Catalyst 6500—1P2Q1T Queuing and Dropping
- Catalyst 6500—1P2Q2T Queuing and Dropping
- Catalyst 6500—1P3Q1T Queuing and Dropping
- Catalyst 6500—1P3Q8T Queuing and Dropping
- Catalyst 6500—1P7Q8T Queuing and Dropping

Catalyst 6500 Queuing and Dropping Overview

While the Catalyst 6500 PFC performs classification, marking, mapping, and policing functions, all queuing and congestion avoidance policies are administered by the Catalyst 6500 linecards. This inevitably leads to per-linecard hardware-specific capabilities and syntax when it comes to configuring queuing and dropping.

As previously discussed in relation to other platforms that support ingress queuing, receive queues are extremely difficult to congest, even in controlled lab environments. This is especially so if access edge policies, as detailed in this chapter, are used on all access layer switches.

Ingress congestion implies that the combined ingress rates of traffic exceed the switch fabric channel speed, and thus would need to be queued simply to gain access to the switching fabric. On newer platforms, such as the Catalyst 6500 Sup720, this means that a combined ingress rate of up to 40 Gbps per slot would be required to create such an event.

However, to obviate such an extreme event, the Catalyst 6500 schedules ingress traffic through the receive queues based on CoS values. In the default configuration, the scheduler assigns all traffic with CoS 5 to the strict-priority queue (if present); in the absence of a strict priority queue, the scheduler assigns all traffic to the standard queues. All other traffic is assigned to the standard queue(s) (with higher CoS values being assigned preference over lower CoS values, wherever supported). Additionally, if a port is configured to trust CoS, then the ingress scheduler implements CoS-value-based receive-queue drop thresholds to avoid congestion in received traffic. Thus, even if the extremely unlikely event of ingress congestion should occur, the default settings for the Catalyst 6500 linecard receive queues are more than adequate to protect VoIP and network control traffic.

Therefore, the focus of this section is on Catalyst 6500 egress/transmit queuing design recommendations.

Catalyst 6500 Transmit Queuing and Dropping Linecard Options

There are **currently** six main transmit queuing/dropping options for Catalyst 6500 linecards:

- 2Q2T—Indicates two standard queues, each with two configurable tail-drop thresholds.
- 1P2Q1T—Indicates one strict-priority queue and two standard queues, each with one configurable WRED-drop threshold (however, each standard queue also has one nonconfigurable tail-drop threshold).

- 1P2Q2T—Indicates one strict-priority queue and two standard queues, each with two configurable WRED-drop thresholds.
- 1P3Q1T—Indicates one strict-priority queue and three standard queues, each with one configurable WRED-drop threshold (however, each standard queue also has one nonconfigurable tail-drop threshold).
- 1P3Q8T—Indicates one strict-priority queue and three standard queues, each with eight configurable WRED-drop thresholds (however, each standard queue also has one nonconfigurable tail-drop threshold).
- 1P7Q8T—Indicates one strict-priority queue and seven standard queues, each with eight configurable WRED-drop thresholds (on 1p7q8t ports, each standard queue also has one nonconfigurable tail-drop threshold).

Almost all Catalyst 6500 linecards support a strict-priority queue and when supported, the switch services traffic in the strict-priority transmit queue before servicing the standard queues. When the switch is servicing a standard queue, after transmitting a packet, it checks for traffic in the strict-priority queue. If the switch detects traffic in the strict-priority queue, it suspends its service of the standard queue and completes service of all traffic in the strict-priority queue before returning to the standard queue.

Additionally, Catalyst 6500 linecards implement CoS-value-based transmit-queue drop thresholds to avoid congestion in transmitted traffic. WRED thresholds can also be defined on certain linecards, where the CoS value of the packet (not the IP Precedence value, although they likely match) determines the WRED weight. WRED parameters include a lower and upper threshold: the low WRED threshold is the queue level where (assigned) traffic begins to be selectively-dropped and the high WRED threshold is the queue level above which all (assigned) traffic is dropped. Furthermore, packets in the queue between the low and high WRED thresholds have an increasing chance of being dropped as the queue fills.

The Transmit Queuing/Dropping capabilities can be returned by using the following commands.

- CatOS:
  - show port capabilities
  - show port qos
  - show qos info
- IOS:
  - show queueing interface

Table 2-5 shows the Catalyst 6500 linecards that are currently available and their respective queuing/dropping structures.

Table 2-5 Catalyst 6500 Linecard Queuing Structures

C2 (xCEF720) Modules	Description	Receive Queue Structure	Transmit Queue Structure	Buffer Size
WS-X6704-10GE	Catalyst 6500 4-port 10 GigabitEthernet Module	1Q8T (8Q8T with DFC3a)	1P7Q8T	16MB per port
WS-X6724-SFP	Catalyst 6500 24-port GigabitEthernet SFP Module	1Q8T; (2Q8T with DFC3a)	1P3Q8T	1MB per port

WS-X6748-GE-TX	Catalyst 6500 48-port 10/100/1000 RJ-45 Module	1Q8T; (2Q8T with DFC3a)	1P3Q8T	1MB per port
WS-X6748-SFP	Catalyst 6500 48-port GigabitEthernet SFP Module	1Q8T; (2Q8T with DFC3a)	1P3Q8T	1MB per port
Classic/CEF256 Ethernet Modules	**Description**	**Receive Queue Structure**	**Transmit Queue Structure**	**Buffer Size**
WS-X6024-10FL-MT	Catalyst 6000 24-port 10BaseFL MT-RJ Module	1Q4T	2Q2T	64KB per port
WS-X6148-RJ21	Catalyst 6500 48-Port 10/100 RJ-21 Module (Upgradable to Voice)	1Q4T	2Q2T	128KB per port
WS-X6148-RJ21V	Catalyst 6500 48-port 10/100 Inline Power RJ-21 Module	1Q4T	2Q2T	128KB per port
WS-X6148-RJ45	Catalyst 6500 48-Port 10/100; RJ-45 Module (Upgradable to Voice)	1Q4T	2Q2T	128KB per port
WS-X6148-RJ45V	Catalyst 6500 48-port 10/100 Inline Power RJ-45 Module	1Q4T	2Q2T	128KB per port
WS-X6148-GE-TX	Catalyst 6500 48-port 10/100/1000 RJ-45 Module	1Q2T	1P2Q2T	1MB per 8 ports
WS-X6148V-GE-TX	Catalyst 6500 48-port 10/100/1000 Inline Power RJ-45 Module	1Q2T	1P2Q2T	1MB per 8 ports
WS-X6316-GE-TX	Catalyst 6000 16-port 1000TX GigabitEthernet RJ-45 Module	1P1Q4T	1P2Q2T	512KB per port
WS-X6324-100FX-MM	Catalyst 6000 24-port 100FX MT-RJ MMF Module (with Enhanced QoS)	1Q4T	2Q2T	128KB per port
WS-X6324-100FX-SM	Catalyst 6000 24-port 100FX MT-RJ SMF Module (with Enhanced QoS)	1Q4T	2Q2T	128KB per port
WS-X6348-RJ-21	Catalyst 6000 48-port 10/100 RJ-21 Module	1Q4T	2Q2T	128KB per port
WS-X6348-RJ21V	Catalyst 6000 48-port 10/100 Inline Power RJ-21 Module	1Q4T	2Q2T	128KB per port
WS-X6348-RJ-45	Catalyst 6500 48-port 10/100 RJ-45 Module (Upgradable to Voice)	1Q4T	2Q2T	128KB per port
WS-X6348-RJ45V	Catalyst 6500 48-port 10/100 Inline Power RJ-45 Module	1Q4T	2Q2T	128KB per port
WS-X6408A-GBIC	Catalyst 6000 8-port GigabitEthernet Module (with Enhanced QoS; Requires GBICs)	1P1Q4T	1P2Q2T	512KB per port
WS-X6416-GBIC	Catalyst 6000 16-port GigabitEthernet Module (Requires GBICs)	1P1Q4T	1P2Q2T	512KB per port
WS-X6416-GE-MT	Catalyst 6000 16-port GigabitEthernet MT-RJ Module	1P1Q4T	1P2Q2T	512KB per port
WS-X6501-10GEX4	1-port 10 GigabitEthernet Module	1P1Q8T	1P2Q1T	64MB per port

WS-X6502-10GE	Catalyst 6500 10 GigabitEthernet Base Module (Requires OIM)	1P1Q8T	1P2Q1T	64MB per port
WS-X6516A-GBIC	Catalyst 6500 16-port GigabitEthernet Module (Fabric-enabled; Requires GBICs)	1P1Q4T	1P2Q2T	1MB per port
WS-X6516-GBIC	Catalyst 6500 16-port GigabitEthernet Module (Fabric-Enabled; Requires GBICs)	1P1Q4T	1P2Q2T	512KB per port
WS-X6516-GE-TX	Catalyst 6500 16-port GigabitEthernet Copper Module; (Crossbar-enabled)	1P1Q4T	1P2Q2T	512KB per port
WS-X6524-100FX-MM	Catalyst 6500 24-port 100FX MT-RJ Module (Fabric-enabled)	1P1Q0T	1P3Q1T	1MB per port
WS-X6548-RJ-21	Catalyst 6500 48-port 10/100 RJ-21 Module (Fabric-enabled)	1P1Q0T	1P3Q1T	1MB per port
WS-X6548-RJ-45	Catalyst 6500 48-port 10/100 RJ-45 Module (Crossbar-enabled)	1P1Q0T	1P3Q1T	1MB per port
WS-X6548V-GE-TX	Catalyst 6500 48-port 10/100/1000 Inline Power RJ-45 Module (Fabric-enabled)	1Q2T	1P2Q2T	1MB per 8 ports
WS-X6548-GE-TX	Catalyst 6500 48-port 10/100/1000 RJ-45 Module (Fabric-enabled)	1Q2T	1P2Q2T	1MB per 8 ports
WS-X6816-GBIC	Catalyst 6500 16-port GigabitEthernet Module (Fabric-Enabled; Requires GBICs)	1P1Q4T	1P2Q2T	512KB per port

Design recommendations for each of these six main Catalyst 6500 queuing structures follow.

Catalyst 6500—2Q2T Queuing and Dropping

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

Linecards that only support 2Q2T queuing models have no provision for priority-queuing. Nonetheless, tuning the Weighted Round-Robin (WRR) weights and the queue sizes can help offset this limitation.

For example, if Q1 is to service Scavenger/Bulk (CoS 1) and Best Effort (CoS 0) traffic, then assigning 30% of the buffer space to the first queue is adequate; the remaining 70% can be assigned to Q2.

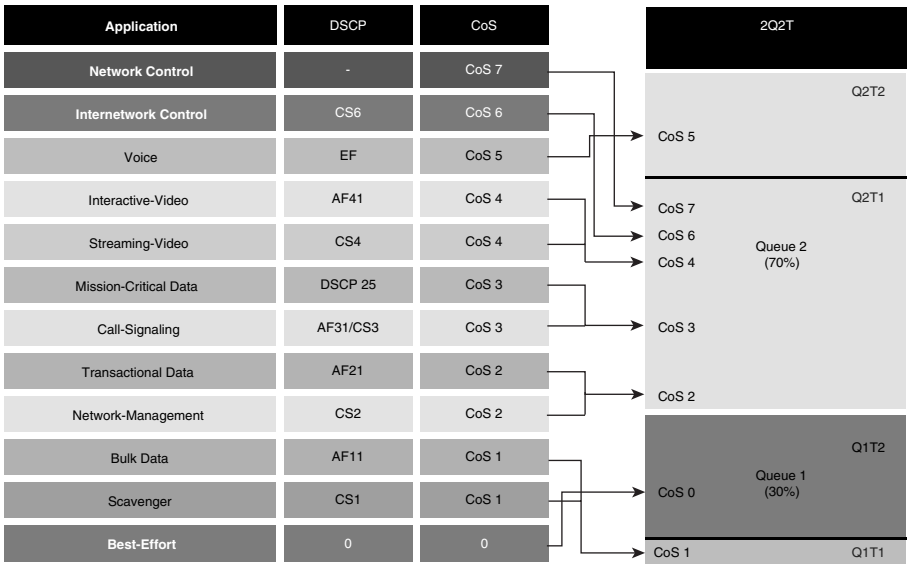
The WRR weights can be set to the same ratio of 30:70 for servicing Q1:Q2.

Since the 2Q2T model supports configurable Tail-Drop thresholds, these can be tuned to provide an additional layer of QoS granularity. For example, the first queue's first threshold can be set at 40% to prevent Scavenger/Bulk traffic from dominating Q1. Similarly, the second queue's first threshold can be set to 80% to always allow some room in the queue for VoIP. The second threshold of each queue should *always* be set to the tail of the queue (100%).

Once the queues and thresholds have been defined as above, then CoS 1 (Scavenger/Bulk) can be assigned to Q1T1; CoS 0 (Best Effort) can be assigned to Q1T2; CoS 2 (Network Management and Transactional Data), CoS 3 (call signaling and Mission-Critical Data), CoS 4 (Interactive and Streaming Video), and CoS 6 and 7 (Internetwork and Network Control) can be assigned to Q21T; and CoS 5 (VoIP) can be assigned to Q2T2.

These 2Q2T queuing recommendations are illustrated in Figure 2-26.

Figure 2-26 Catalyst 6500 2Q2T Queuing Model



Configuration

The Catalyst 6500 CatOS configurations to configure 2Q2T queuing recommendations are shown below.

Example 2-64 Catalyst 6500 CatOS—2Q2T Queuing Example

```
CAT6500-PFC2-CATOS> (enable) set qos txq-ratio 2q2t 30 70
    ! Sets the buffer allocations to 30% for Q1 and 70% for Q2
CAT6500-PFC2-CATOS> (enable) set qos wrr 2q2t 30 70
    ! Sets the WRR weights for 30:70 (Q1:Q2) bandwidth servicing
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos drop-threshold 2q2t tx queue 1 40 100
    ! Sets Q1T1 to 5% to limit Scavenger/Bulk from dominating Q1
CAT6500-PFC2-CATOS> (enable) set qos drop-threshold 2q2t tx queue 2 80 100
    ! Sets Q2T1 to 80% to always have room in Q2 for VoIP
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos map 2q2t tx 1 1 cos 1
    ! Assigns Scavenger/Bulk to Q1T1
CAT6500-PFC2-CATOS> (enable) set qos map 2q2t tx 1 2 cos 0
    ! Assigns Best Effort to Q1T2
CAT6500-PFC2-CATOS> (enable) set qos map 2q2t tx 2 1 cos 2,3,4,6,7
    ! Assigns CoS 2,3,4,6 and 7 to Q2T1
CAT6500-PFC2-CATOS> (enable) set qos map 2q2t tx 2 2 cos 5
    ! Assigns VoIP to Q2T2
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Commands:

- show qos info config 2q2t tx
- show qos info runtime
- show qos statistics

Catalyst 6500 CatOS QoS Verification Command: show qos info config 2q2t tx

The Catalyst 6500 CatOS **show qos info config 2q2t tx** verification command displays the queuing and dropping parameters for 2Q2T linecards.

In the example below, CoS 1 is assigned to Q1T1; CoS 0 is assigned to Q1T2; CoS values 2,3,4,6 and 7 are assigned to Q2T1 and CoS 5 is assigned to Q2T2. The first thresholds are set to 40% and 80%of their respective queues, and the second thresholds set to the tail of the queue. The size ratio has been allocated 30% for Q1 and 70% for Q2 and the WRR weights are set to 30:70 to service Q1 and Q2, respectively.

Example 2-65 Show QoS Info Config 2Q2T Tx Verification for a Catalyst 6500-CatOS Switch

```
CAT6500-PFC2-CATOS> (enable) show qos info config 2q2t tx
QoS setting in NVRAM for 2q2t transmit:
QoS is enabled
Queue and Threshold Mapping for 2q2t (tx):
Queue Threshold CoS
----- --------- ---------------
1     1         1
1     2         0
2     1         2 3 4 6 7
2     2         5
Tx drop thresholds:
Queue #  Thresholds - percentage
-------  ------------------------------------
1        40%  100%
2        80% 100%
Tx WRED thresholds:
WRED feature is not supported for this port type.
Tx queue size ratio:
Queue #  Sizes - percentage
-------  ------------------------------------
1        30%
2        70%
Tx WRR Configuration of ports with 2q2t:
Queue #  Ratios
-------  ------------------------------------
1        30
2        70
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Command: show qos info runtime

The Catalyst 6500 CatOS **show qos info runtime** verification command reports similar information as the **show qos info config** command, but displays the runtime information (committed to the PFC and linecard) as opposed to only the configured information.

In the example below, CoS 1 is assigned to Q1T1; CoS 0 is assigned to Q1T2; CoS values 2,3,4,6 and 7 are assigned to Q2T1 and CoS 5 is assigned to Q2T2. The first thresholds are set to 40% and 80% of their respective queues, and the second thresholds set to the tail of the queue. The size ratio has been allocated 30% for Q1 and 70% for Q2 and the WRR weights are set to 30:70 to service Q1 and Q2, respectively.

Example 2-66 Show QoS Info Runtime Verification for a Catalyst 6500-CatOS Switch

```
CAT6500-PFC3-CATOS> (enable) show qos info runtime 3/1
Run time setting of QoS:
QoS is enabled
Policy Source of port 3/1: Local
Tx port type of port 3/1 : 2q2t
Rx port type of port 3/1 : 1q4t
Interface type: port-based
ACL attached:
The qos trust type is set to untrusted.
Default CoS = 0
Queue and Threshold Mapping for 2q2t (tx):
Queue Threshold CoS
----- --------- ---------------
1     1         1
1     2         0
2     1         2 3 4 6 7
2     2         5
Queue and Threshold Mapping for 1q4t (rx):
All packets are mapped to a single queue.
Rx drop thresholds:
Rx drop thresholds are disabled.
Tx drop thresholds:
Queue #  Thresholds - percentage (* abs values)
-------  ------------------------------------
1        40% (6144 bytes) 100% (15360 bytes)
2        80% (28672 bytes) 100% (35840 bytes)
Rx WRED thresholds:
Rx WRED feature is not supported for this port type.
Tx WRED thresholds:
WRED feature is not supported for this port type.
Tx queue size ratio:
Queue #  Sizes - percentage (* abs values)
-------  ------------------------------------
1        30% (17408 bytes)
2        70% (37888 bytes)
Rx queue size ratio:
Rx queue size-ratio feature is not supported for this port type.
Tx WRR Configuration of ports with speed 10Mbps:
Queue #  Ratios (* abs values)
-------  ------------------------------------
1        30 (7648 bytes)
2        70 (17840 bytes)
(*) Runtime information may differ from user configured setting due to hardware
granularity.
CAT6500-PFC3-CATOS> (enable)
```

The Catalyst 6500 IOS configurations to configure 2Q2T queuing recommendations are shown below.

Example 2-67 Catalyst 6500 IOS—2Q2T Queuing Example

```
CAT6500-PFC3-IOS(config)# interface range FastEthernet6/1 - 48
CAT6500-PFC3-IOS(config-if)# wrr-queue queue-limit 30 70
    ! Sets the buffer allocations to 30% for Q1 and 70% for Q2
CAT6500-PFC3-IOS(config-if)# wrr-queue bandwidth 30 70
    ! Sets the WRR weights for 30:70 (Q1:Q2) bandwidth servicing
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue threshold 1 40 100
    ! Sets Q1T1 to 5% to limit Scavenger/Bulk from dominating Q1
CAT6500-PFC3-IOS(config-if)# wrr-queue threshold 2 80 100
    ! Sets Q2T1 to 80% to always have room in Q2 for VoIP
CAT6500-PFC3-IOS(config-if)#
```

```
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 1 1 1
    ! Assigns Scavenger/Bulk to Q1T1
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 1 2 0
    ! Assigns Best Effort to Q1T2
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 2 1 2 3 4 6 7
    ! Assigns CoS 2,3,4,6 and 7 to Q2T1
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 2 2 5
    ! Assigns VoIP to Q2T2
CAT6500-PFC3-IOS(config-if)#end
CAT6500-PFC3-IOS#
```

Catalyst 6500 MLS QoS Verification Commands:

- show queueing interface

Catalyst 6500 IOS QoS Verification Command: show queueing interface

The Catalyst 6500 IOS **show queueing interface** verification command displays the queuing parameters for a given interface (according to the linecard's capabilities).

In the example below, the linecard has 2Q2T transmit queuing. The WRR scheduling weights are set to 30:70 to service Q1 and Q2, respectively. The transmit queue size ratios have been allocated 30% for Q1 and 70 percent for Q2. The first queue's tail-drop thresholds are set to 40% and 100%, while the second queue's tail-drop thresholds are set to 80% and 100%. CoS 1 is assigned to Q1T1; CoS 0 is assigned to Q1T2; CoS values 2,3,4,6 and 7 are assigned to Q2T1 and CoS 5 is assigned to Q2T2.

Example 2-68 Show Queueing Interface Verification for a Catalyst 6500-IOS Switch

```
CAT6500-PFC3-IOS#show queueing interface FastEthernet6/1
Interface FastEthernet6/1 queueing strategy:  Weighted Round-Robin
  Port QoS is enabled
  Port is untrusted
  Extend trust state: not trusted [COS = 0]
  Default COS is 0
    Queueing Mode In Tx direction: mode-cos
    Transmit queues [type = 2q2t]:
    Queue Id    Scheduling  Num of thresholds
    -----------------------------------------
       1         WRR low             2
       2         WRR high            2

    WRR bandwidth ratios:   30[queue 1]  70[queue 2]
    queue-limit ratios:     30[queue 1]  70[queue 2]

    queue tail-drop-thresholds
    --------------------------
    1     40[1] 100[2]
    2     80[1] 100[2]

    queue thresh cos-map
    ---------------------------------------
    1     1      1
    1     2      0
    2     1      2 3 4 6 7
    2     2      5

    <output truncated>

CAT6500-PFC3-IOS#
```

Catalyst 6500—1P2Q1T Queuing and Dropping

This section includes the following topics:

- Configuration
- Catalyst MLS QoS Verification Commands

The 1P2Q1T queuing model builds on the previous 2Q2T model, bringing with it the advantages of strict-priority queuing (for VoIP) as well as a tunable WRED (not Tail-Drop) threshold per queue.

The term "1P2Q1T" is a bit of a misnomer in the CatOS version of this queuing structure because in CatOS there are actually two thresholds per queue: the tunable WRED threshold and the non-configurable tail-of-the-queue (100%) tail-drop threshold.

Under such a model, buffer space can be allocated as follows: 30% for Scavenger/Bulk with Best Effort queue (Q1), 40% for Q2, and 30% for the PQ (Q3)

The WRR weights for Q1 and Q2 (for dividing the remaining bandwidth, after the priority queue has been fully serviced) can be set to 30:70 respectively for Q1:Q2.

Under the 1P2Q1T model, each queue's WRED threshold is defined with a lower and upper limit. For example, the WRED threshold 40:80 indicates that packets assigned to this WRED threshold will *begin* being randomly dropped when the queue fills to 40 percent *and* that these packets will be tail-dropped if the queue fills beyond 80 percent.

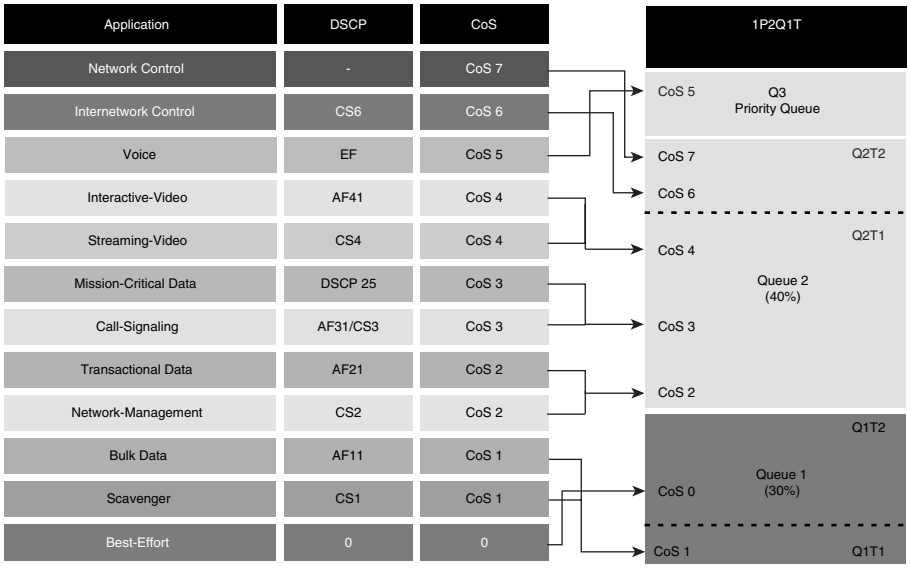
Furthermore, in CatOS within the 1P2Q1T queuing structure, each CoS value can be assigned to queue *and* a WRED threshold or just to a queue. When assigned to a queue (only), then the CoS value is limited only by the tail of the queue (in other words, it is assigned to the queue with a tail-drop threshold of 100%).

Thus (in CatOS), the tunable WRED threshold for Q1 can be set to 40:80, meaning that Scavenger/Bulk Data will be WRED-dropped if Q1 fills to 40 percent and will be tail-dropped if Q1 fills past 80 percent of capacity. This prevents Scavenger/Bulk Data from drowning out Best-Effort traffic in Q1. The WRED threshold for Q2 can be set to 70:80 to provide congestion avoidance for all applications assigned to it and to ensure that there will always be room in the queue to service Network and Internetwork Control traffic.

Therefore, once the queues and thresholds have been defined as above, then CoS 1 (Scavenger/Bulk) can be assigned to Q1T1; CoS 0 (Best Effort) to Q1-only (tail); CoS 2 (Network Management and Transactional Data), CoS 3 (call signaling and Mission-Critical Data) and CoS 4 (Interactive and Streaming Video) can be assigned to Q2T1; CoS 6 and 7 (Internetwork and Network Control) can be assigned to Q2-only (tail); CoS 5 (VoIP) can be assigned to Q3 (the PQ).

These 1P2Q1T queuing recommendations are illustrated in Figure 2-27.

Figure 2-27 Catalyst 6500 1P2Q1T Queuing Model (CatOS Supports 1P2Q2T)



Configuration

The Catalyst 6500 CatOS configurations to configure 1P2Q1T queuing recommendations are shown below.

Example 2-69 Catalyst 6500 CatOS—1P2Q1T Queuing Example (technically 1P2Q2T)

```
CAT6500-PFC2-CATOS> (enable) set qos txq-ratio 1p2q1t 30 40 30
! Sets the buffer allocations to 30% for Q1, 40% for Q2, 30% for Q3 (PQ)
CAT6500-PFC2-CATOS> (enable) set qos wrr 1p2q1t 30 70
! Sets the WRR weights for 30:70 (Q1:Q2) bandwidth servicing
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos wred 1p2q1t tx queue 1 40:80
! Sets Q1 WRED Threshold to 40:80 to limit Scavenger/Bulk from dominating Q1
CAT6500-PFC2-CATOS> (enable) set qos wred 1p2q1t tx queue 2 70:80
! Sets Q2 WRED Threshold to 70:80 to force room for Network Control traffic
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q1t tx 1 1 cos 1
    ! Assigns Scavenger/Bulk to Q1 WRED Threshold
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q1t tx 1 cos 0
    ! Assigns Best Effort to Q1 tail (100%) threshold
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q1t tx 2 1 cos 2,3,4
    ! Assigns CoS 2,3,4 to Q2 WRED Threshold
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q1t tx 2 cos 6,7
    ! Assigns Network/Internetwork Control to Q2 tail (100%) threshold
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q1t tx 3 cos 5
    ! Assigns VoIP to PQ (Q3)
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Commands:

- show qos info config 1p2q1t tx
- show qos info runtime
- show qos statistics

Note The Catalyst 6500 CatOS **show qos info** verification commands are reasonably similar for each queuing structure and as such (to minimize redundancy) are not detailed for each queuing model example.

In IOS, for any 1P*x*Q*y*T queuing structure, setting the size of the priority queue is not supported. The only exception to this is within the 1P2Q2T structure, where the priority queue (Q3) is indirectly set to equal Q2's size. Therefore, in all examples of Catalyst 6500 IOS queuing structure configurations to follow (that support a PQ) only the sizes of the standard queues are being set.

Furthermore, specific to the 1P2Q1T queuing structure, CoS values cannot be mapped to the tail of the queue, as in CatOS. CoS values can be mapped only to the single WRED threshold for each queue. Therefore, the 1P2Q1T queuing and dropping recommendation requires some slight alterations for Cisco IOS. These include changing Q1T1's WRED threshold to 80:100 and, likewise, changing Q2T1's WRED threshold to 80:100.

The syntax-logic for setting WRED thresholds in IOS is different from CatOS. In CatOS, minimum and maximum WRED thresholds were set on the same line; in IOS, minimum and maximum WRED thresholds are set on different lines.

After these WRED thresholds have been altered, then CoS 1 (Scavenger/Bulk) and CoS 0 (Best Effort) can be assigned to Q1T1; CoS 2 (Network Management and Transactional Data), CoS 3 (call signaling and Mission-Critical Data), CoS 4 (Interactive and Streaming Video) and CoS 6 and 7 (Internetwork and Network Control) can be assigned to Q2T1; CoS 5 (VoIP) can be assigned to Q3 (the PQ).

The Catalyst 6500 IOS configurations to configure 1P2Q1T queuing recommendations are shown below.

Example 2-70 Catalyst 6500 IOS—1P2Q1T Queuing Example

```
CAT6500-PFC3-IOS(config)#interface TenGigabitEthernet1/1
CAT6500-PFC3-IOS(config-if)# wrr-queue queue-limit 30 40
! Sets the buffer allocations to 30% for Q1 and 40% for Q2
CAT6500-PFC3-IOS(config-if)# wrr-queue bandwidth 30 70
! Sets the WRR weights for 30:70 (Q1:Q2) bandwidth servicing
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect min-threshold 1 80
    ! Sets Min WRED Threshold for Q1T1 to 80%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect max-threshold 1 100
    ! Sets Max WRED Threshold for Q1T1 to 100%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect min-threshold 2 80
    ! Sets Min WRED Threshold for Q2T1 to 80%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect max-threshold 2 100
    ! Sets Max WRED Threshold for Q2T1 to 100%
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 1 1 1 0
    ! Assigns Scavenger/Bulk and Best Effort to Q1 WRED Threshold 1
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 2 1 2 3 4 6 7
    ! Assigns CoS 2,3,4,6 and 7 to Q2 WRED Threshold 1
CAT6500-PFC3-IOS(config-if)# priority-queue cos-map 1 5
    ! Assigns VoIP to PQ (Q3)
CAT6500-PFC3-IOS(config-if)#end
CAT6500-PFC3-IOS(config-if)#
```

Catalyst 6500 MLS QoS Verification Commands:

- show queueing interface

Catalyst 6500—1P2Q2T Queuing and Dropping

This section includes the following topics:

- Recommendations
- Configuration
- Catalyst MLS QoS Verification Commands

Recommendations

The 1P2Q2T queuing model is essentially identical to the 1P2Q1T model, except that it supports two configurable WRED thresholds per queue. Under this model, CoS values cannot be mapped to the tail of the queue, as with the 1P2Q1T model, and so there are (as the name correctly implies this time) two effective thresholds per queue.

Under a 1P2Q2T model, buffer space can be allocated as follows: 30% for Q1 (the Scavenger/Bulk with Best Effort queue), 40% for Q2 (the preferential queue), and 30% for the Q3 (the priority queue).

The WRR weights for Q1 and Q2 (for dividing the remaining bandwidth, after the priority queue has been fully serviced) remain at 30:70 respectively for Q1:Q2.

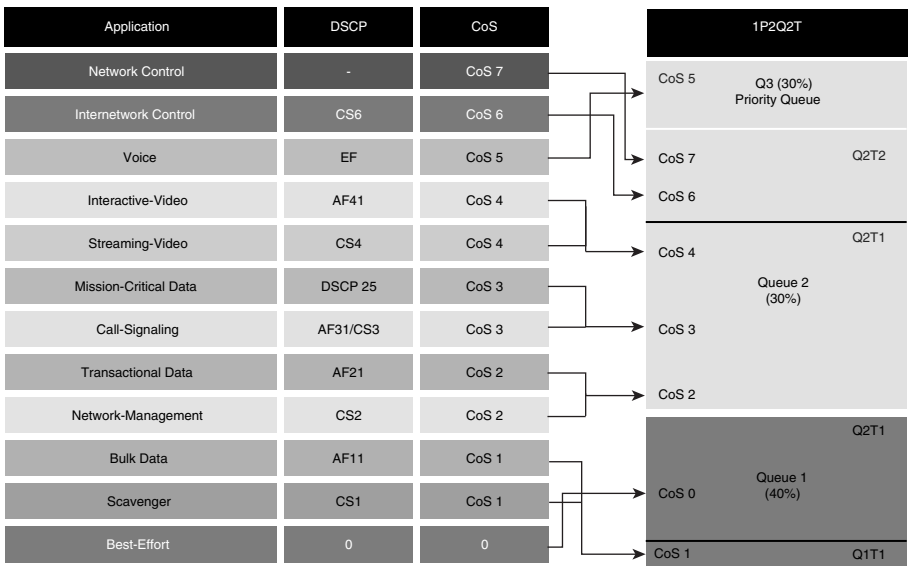
Under the 1P2Q2T model, each WRED threshold is defined with a lower and upper limit. Therefore, the first WRED threshold for Q1 can be set to 40:80, so that Scavenger/Bulk Data traffic can be WRED-dropped if Q1 hits 40 percent and can be tail-dropped if Q1 exceeds 80 percent of its capacity (this prevents Scavenger/Bulk Data from drowning out Best-Effort traffic in Q1). The second WRED threshold for Q1 can be set to 80:100 to provide congestion avoidance for Best-Effort traffic.

Similarly, the first WRED threshold of Q2 can be set to 70:80, and the second can be set to 80:100. In this manner, congestion avoidance will be provided for all traffic types in Q2, and there will always be room in the queue to service Network and Internetwork Control traffic.

Therefore, once the queues have been defined as above, then CoS 1 (Scavenger/Bulk) can be assigned to Q1T1; CoS 0 (Best Effort) to Q1T2; CoS 2 (Network Management and Transactional Data), CoS 3 (call signaling and Mission-Critical Data) and CoS 4 (Interactive and Streaming Video) can be assigned to Q2T1; CoS 6 and 7 (Internetwork and Network Control) can be assigned to Q2T2; CoS 5 (VoIP) can be assigned to Q3T1 (the PQ).

These 1P2Q2T queuing recommendations are illustrated in Figure 2-28.

Figure 2-28 Catalyst 6500 1P2Q2T Queuing Model



Configuration

The Catalyst 6500 CatOS configurations to configure 1P2Q1T queuing recommendations are shown below.

Example 2-71 Catalyst 6500 CatOS—1P2Q2T Queuing Example

```
CAT6500-PFC2-CATOS> (enable) set qos txq-ratio 1p2q2t 30 40 30
! Sets the buffer allocations to 30% for Q1, 40% for Q2, 30% for Q3 (PQ)
CAT6500-PFC2-CATOS> (enable) set qos wrr 1p2q2t 30 70
! Sets the WRR weights for 30:70 (Q1:Q2) bandwidth servicing
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos wred 1p2q2t tx queue 1 40:80 80:100
! Sets Q1 WRED T1 to 40:80 to limit Scavenger/Bulk from dominating Q1
! Sets Q1 WRED T2 to 80:100 to provide congestion-avoidance for Best Effort
CAT6500-PFC2-CATOS> (enable) set qos wred 1p2q2t tx queue 2 70:80 80:100
! Sets Q2 WRED T1 to 70:80 to provide congestion-avoidance
! Sets Q2 WRED T2 to 80:100 to force room for Network Control traffic
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q2t tx 1 1 cos 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q2t tx 1 2 cos 0
! Assigns Best Effort to Q1 WRED Threshold 2
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q2t tx 2 1 cos 2,3,4
! Assigns CoS 2,3,4 to Q2 WRED Threshold 1
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q2t tx 2 2 cos 6,7
! Assigns Network/Internetwork Control to Q2 WRED Threshold 2
CAT6500-PFC2-CATOS> (enable) set qos map 1p2q2t tx 3 1 cos 5
! Assigns VoIP to PQ
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Commands:

- show qos info config 1p2q2t tx
- show qos info runtime
- show qos statistics

The compatible Catalyst 6500 IOS configurations to configure 1P2Q1T queuing recommendations are shown below. Notice that the buffer allocation for the PQ (Q3) is not configurable, but is, by default (for the 1P2Q2T queuing structure only) set to equal the size defined for Q2. Therefore, the queue size ratios have been slightly altered from the CatOS version of this queuing model to take this default IOS behavior into account; specifically Q1 is set to 40% and Q2 is set to 30% (which indirectly sets Q3 to match, at 30%).

The Catalyst 6500 IOS configurations to configure 1P2Q2T queuing recommendations are shown below.

Example 2-72 Catalyst 6500 IOS—1P2Q2T Queuing

```
CAT6500-PFC3-IOS(config)#interface range GigabitEthernet4/1 - 8
CAT6500-PFC3(config-if-range)# wrr-queue queue-limit 40 30
! Sets the buffer allocations to 40% for Q1 and 30% for Q2
! Indirectly sets PQ (Q3) size to equal Q2 (which is set to 30%)
CAT6500-PFC3(config-if-range)# wrr-queue bandwidth 30 70
! Sets the WRR weights for 30:70 (Q1:Q2) bandwidth servicing
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)# wrr-queue random-detect min-threshold 1 40 80
    ! Sets Min WRED Thresholds for Q1T1 and Q1T2 to 40 and 80, respectively
CAT6500-PFC3(config-if-range)# wrr-queue random-detect max-threshold 1 80 100
    ! Sets Max WRED Thresholds for Q1T1 and Q1T2 to 80 and 100, respectively
CAT6500-PFC3(config-if-range)#
```

```
CAT6500-PFC3(config-if-range)# wrr-queue random-detect min-threshold 2 70 80
    ! Sets Min WRED Thresholds for Q2T1 and Q2T2 to 70 and 80, respectively
CAT6500-PFC3(config-if-range)# wrr-queue random-detect max-threshold 2 80 100
    ! Sets Max WRED Thresholds for Q2T1 and Q2T2 to 80 and 100, respectively
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 1 1 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 1 2 0
! Assigns Best Effort to Q1 WRED Threshold 2
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 2 1 2 3 4
! Assigns CoS 2,3,4 to Q2 WRED Threshold 1
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 2 2 6 7
! Assigns Network/Internetwork Control to Q2 WRED Threshold 2
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)# priority-queue cos-map 1 5
! Assigns VoIP to PQ
CAT6500-PFC3(config-if-range)#end
CAT6500-PFC3-IOS#
```

Catalyst 6500 MLS QoS Verification Commands:

- show queueing interface

Catalyst 6500—1P3Q1T Queuing and Dropping

This section includes the following topics:

- Recommendations
- Configuration
- Catalyst MLS QoS Verification Commands

Recommendations

The 1P3Q1T queuing structure is identical to the 1P2Q1T structure, except that an additional standard queue has been added to it and it does not support tuning the Transmit Size Ratios. Under this model, Q4 is the strict-priority queue.

The WRR weights for the standard queues (Q1, Q2, Q3), for dividing the remaining bandwidth, after the priority queue has been fully serviced, can be set to 5:25:70 respectively for Q1:Q2:Q3.

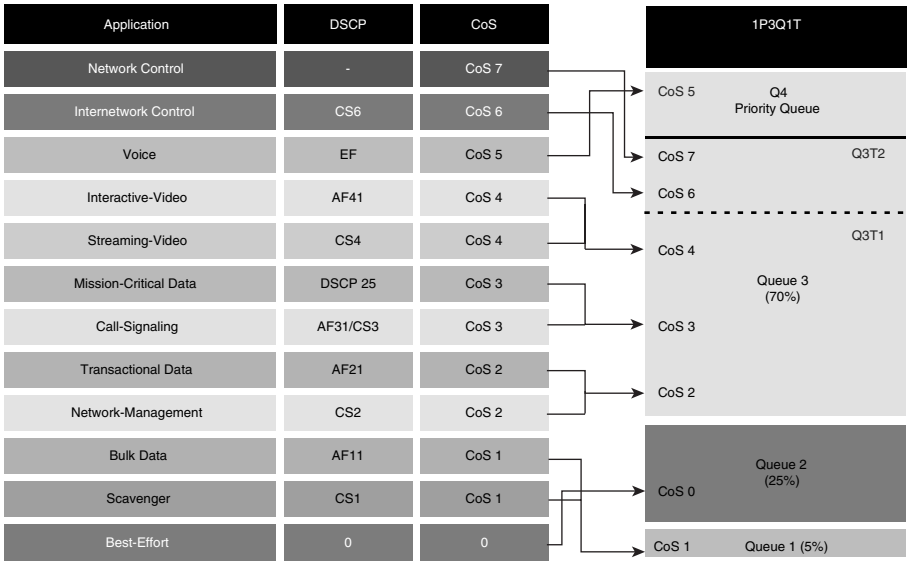
In CatOS, within the 1P3T1T queuing structure each CoS value can be assigned to queue *and* a WRED threshold or just to a queue. When assigned to a queue (only), then the CoS value is limited only by the tail of the queue (in other words, it is assigned to the queue with a tail-drop threshold of 100%).

Thus, the tunable WRED threshold for Q1 can be set to 80:100 to provide congestion avoidance for Scavenger/Bulk Data traffic. The WRED threshold for Q2 similarly can be set to 80:100 to provide congestion avoidance on all Best-Effort flows. The WRED threshold for Q3 can be set to 70:80, to provide congestion avoidance for all applications assigned to it and to ensure that there will always be room in the Q3 to service Network and Internetwork Control traffic.

Therefore, once the queues and thresholds have been defined as above, then CoS 1 (Scavenger/Bulk) can be assigned to Q1T1; CoS 0 (Best Effort) to Q2T1; CoS 2 (Network Management and Transactional Data), CoS 3 (call signaling and Mission-Critical Data) and CoS 4 (Interactive and Streaming Video) can be assigned to Q3T1; CoS 6 and 7 (Internetwork and Network Control) can be assigned to Q3 (tail); CoS 5 (VoIP) can be assigned to Q4 (the PQ).

These 1P3Q1T queuing recommendations are illustrated in Figure 2-29.

Figure 2-29 Catalyst 6500 1P3Q1T Queuing Model (CatOS supports 1P3Q2T)



Configuration

The Catalyst 6500 CatOS configurations to configure 1P3Q1T queuing recommendations are shown below.

Example 2-73 Catalyst 6500 CatOS—1P3Q1T Queuing Example (technically 1P3Q2T)

```
CAT6500-PFC2-CATOS> (enable) set qos wrr 1p3q1t 5 25 70
! Sets the WRR weights for 5:25:70 (Q1:Q2:Q3) bandwidth servicing
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos wred 1p3q1t tx queue 1 80:100
    ! Sets Q1 WRED T1 to 80:100 to provide congestion-avoidance for Scavenger/Bulk
CAT6500-PFC2-CATOS> (enable) set qos wred 1p3q1t tx queue 2 80:100
    ! Sets Q2 WRED T1 to 80:100 to provide congestion-avoidance for Best Effort
CAT6500-PFC2-CATOS> (enable) set qos wred 1p3q1t tx queue 3 70:80
! Sets Q3 WRED T1 to 70:80 to provide congestion-avoidance for CoS 2,3,4
! and to force room (via tail-drop) for Network Control traffic
CAT6500-PFC2-CATOS> (enable)
CAT6500-PFC2-CATOS> (enable) set qos map 1p3q1t tx 1 1 cos 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1 (80:100)
CAT6500-PFC2-CATOS> (enable) set qos map 1p3q1t tx 2 1 cos 0
! Assigns Best Effort to Q2 WRED Threshold 1 (80:100)
CAT6500-PFC2-CATOS> (enable) set qos map 1p3q1t tx 3 1 cos 2,3,4
! Assigns CoS 2,3,4 to Q3 WRED Threshold 1 (70:80)
CAT6500-PFC2-CATOS> (enable) set qos map 1p3q1t tx 3 cos 6,7
! Assigns Network/Internetwork Control to Q3 Tail (100%)
CAT6500-PFC2-CATOS> (enable) set qos map 1p3q1t tx 4 cos 5
! Assigns VoIP to PQ (Q4)
CAT6500-PFC2-CATOS> (enable)
```

Catalyst 6500 CatOS QoS Verification Commands:

- show qos info config 1p3q1t tx
- show qos info runtime
- show qos statistics

In IOS, the 1P3Q1T, 1P3Q8T, and 1P7Q8T queuing structures can be configured to use tail-drop or WRED. By default, WRED is disabled. Therefore, it is good practice to always explicitly enable WRED on a queue before setting WRED thresholds for these queuing structures.

Additionally, in Cisco IOS, the 1P3Q1T queuing structure does not support mapping CoS values to the tail of the queue (only to the single WRED threshold). Therefore, the queuing recommendation requires slight alterations for Cisco IOS: changing all three WRED thresholds to 80:100 and mapping CoS values 2, 3, 4, 6, and 7 to Q3T1.

The Catalyst 6500 IOS configurations to configure 1P3Q1T queuing recommendations are shown below.

Example 2-74 Catalyst 6500 IOS—1P3Q1T Queuing Example

```
CAT6500-PFC3-IOS(config)# interface range FastEthernet3/1 - 48
CAT6500-PFC3-IOS(config-if)# wrr-queue bandwidth 5 25 70
! Sets the WRR weights for 5:25:70 (Q1:Q2:Q3) bandwidth servicing
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 1
    ! Enables WRED on Q1
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 2
! Enables WRED on Q2
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 3
! Enables WRED on Q3
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect min-threshold 1 80
    ! Sets Min WRED Threshold for Q1T1 to 80%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect max-threshold 1 100
    ! Sets Max WRED Threshold for Q1T1 to 100%
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect min-threshold 2 80
    ! Sets Min WRED Threshold for Q2T1 to 80%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect max-threshold 2 100
    ! Sets Max WRED Threshold for Q2T1 to 100%
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect min-threshold 3 80
    ! Sets Min WRED Threshold for Q3T1 to 80%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect max-threshold 3 100
    ! Sets Max WRED Threshold for Q3T1 to 100%
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 1 1 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1 (80:100)
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 2 1 0
! Assigns Best Effort to Q2 WRED Threshold 1 (80:100)
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 3 1 2 3 4 6 7
! Assigns CoS 2,3,4,6 and 7 to Q3 WRED Threshold 1 (80:100)
CAT6500-PFC3-IOS(config-if)# priority-queue cos-map 1 5
! Assigns VoIP to PQ (Q4)
CAT6500-PFC3-IOS(config-if)#end
CAT6500-PFC3-IOS#
```

Catalyst 6500 MLS QoS Verification Commands:

- show queueing interface

Catalyst 6500—1P3Q8T Queuing and Dropping

The 1P3Q8T queuing structure is identical to the 1P3Q1T structure, except it has eight tunable WRED thresholds per queue (instead of one) and it also supports tuning the Transmit Size Ratios. Under this model, Q4 is the strict-priority queue.

Under a 1P3Q8T model, buffer space can be allocated as follows: 5% for the Scavenger/Bulk queue (Q1), 25% for the Best Effort queue (Q2), 40% for the preferential queue (Q3), and 30% for the strict priority queue (Q4).

The WRR weights for the standard queues (Q1, Q2, Q3), for dividing the remaining bandwidth, after the priority queue has been fully serviced, can be set to 5:25:70 respectively for Q1:Q2:Q3.

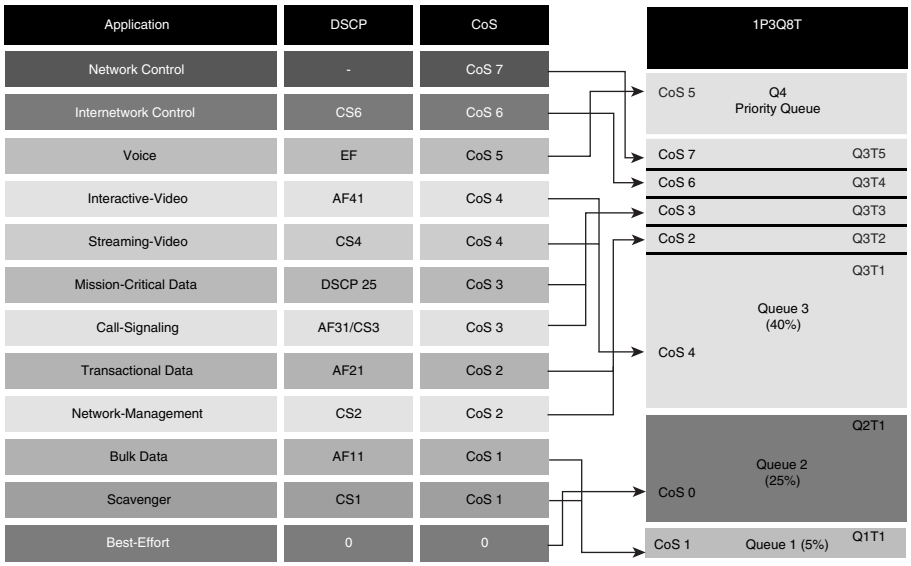
The tunable WRED threshold for Q1 can be set to 80:100 to provide congestion avoidance to Scavenger/Bulk Data traffic. The WRED threshold for Q2 similarly can be set to 80:100 to provide congestion avoidance on all Best-Effort flows.

The 1P3Q8T queuing structure's support for up to eight WRED thresholds per queue allows for additional QoS granularity for the applications sharing Q3. Because only five discrete CoS values are sharing this queue, only five of eight thresholds need to be defined for subqueue QoS. For example, Q3T1 could be set to 50:60, Q3T2 could be set to 60:70, Q3T3 could be set to 70:80, Q3T4 could be set to 80:90, and Q3T5 could be set to 90:100.

Therefore, once the queues and thresholds have been defined as above, CoS 1 (Scavenger/Bulk) can be assigned to Q1T1; CoS 0 (Best Effort) to Q2T1; CoS 4 (Interactive and Streaming Video) can be assigned to Q3T1; CoS 2 (Network Management and Transactional Data) could be assigned to Q3T2; CoS 3 (call signaling and Mission-Critical Data) could be assigned to Q3T3; CoS 6 (Internetwork Control) could be assigned to Q3T4; CoS 7 (Internetwork and Network Control) can be assigned to Q3T5; CoS 5 (VoIP) can be assigned to Q4 (the PQ).

These 1P3Q8T queuing recommendations are illustrated in Figure 2-30.

Figure 2-30 Catalyst 6500 1P3Q8T Queuing Model



The Catalyst 6500 (PFC3) CatOS configurations to configure 1P3Q8T queuing recommendations are shown below.

Example 2-75 Catalyst 6500 (PFC3) CatOS—1P3Q8T Queuing Example

```
CAT6500-PFC3-CATOS> (enable) set qos txq-ratio 1p3q8t 5 25 40 30
! Allocates 5% for Q1, 25% for Q2, 40% for Q3 and 30% for Q4 (PQ)
CAT6500-PFC3-CATOS> (enable) set qos wrr 1p3q8t 5 25 70
! Sets the WRR weights for 5:25:70 (Q1:Q2:Q3) bandwidth servicing
CAT6500-PFC3-CATOS> (enable)
```

```
CAT6500-PFC3-CATOS> (enable) set qos wred 1p3q8t tx queue 1 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q1 WRED T1 to 80:100 and all other Q1 WRED Thresholds to 100:100
CAT6500-PFC3-CATOS> (enable) set qos wred 1p3q8t tx queue 2 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q2 WRED T1 to 80:100 and all other Q2 WRED Thresholds to 100:100
CAT6500-PFC3-CATOS> (enable) set qos wred 1p3q8t tx queue 3 50:60 60:70 70:80
80:90 90:100 100:100 100:100 100:100
! Sets Q3 WRED T1 to 50:60, Q3T2 to 60:70, Q3T3 to 70:80,
! Q3T4 to 80:90, Q3T5 to 90:100
! All other Q3 WRED Thresholds are set to 100:100
CAT6500-PFC3-CATOS> (enable)
CAT6500-PFC3-CATOS> (enable) set qos map 1p3q8t tx 1 1 cos 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
CAT6500-PFC3-CATOS> (enable) set qos map 1p3q8t tx 2 1 cos 0
! Assigns Best Effort to Q2 WRED Threshold 1
CAT6500-PFC3-CATOS> (enable) set qos map 1p3q8t tx 3 1 cos 4
    ! Assigns Video to Q3 WRED Threshold 1
CAT6500-PFC3-CATOS> (enable) set qos map 1p3q8t tx 3 2 cos 2
    ! Assigns Net-Mgmt and Transactional Data to Q3 WRED T2
CAT6500-PFC3-CATOS> (enable) set qos map 1p3q8t tx 3 3 cos 3
! Assigns call signaling and Mission-Critical Data to Q3 WRED T3
CAT6500-PFC3-CATOS> (enable) set qos map 1p3q8t tx 3 4 cos 6
! Assigns Internetwork-Control (IP Routing) to Q3 WRED T4
CAT6500-PFC3-CATOS> (enable) set qos map 1p3q8t tx 3 5 cos 7
! Assigns Network-Control (Spanning Tree) to Q3 WRED T5
CAT6500-PFC3-CATOS> (enable) set qos map 1p3q8t tx 4 cos 5
    ! Assigns VoIP to the PQ (Q4)
CAT6500-PFC3-CATOS> (enable)
```

Catalyst 6500 (PFC3) CatOS QoS Verification Commands:

- show qos info config 1p3q8t tx
- show qos info runtime
- show qos statistics

The Catalyst 6500 (PFC3) IOS configurations to configure 1P3Q8T queuing recommendations are shown below.

Example 2-76 Catalyst 6500 IOS—1P3Q8T Queuing Example

```
CAT6500-PFC3-IOS(config)# interface range GigabitEthernet1/1 - 48
CAT6500-PFC3-IOS(config-if)# wrr-queue queue-limit 5 25 40
! Allocates 5% for Q1, 25% for Q2 and 40% for Q3
CAT6500-PFC3-IOS(config-if)# wrr-queue bandwidth 5 25 70
! Sets the WRR weights for 5:25:70 (Q1:Q2:Q3) bandwidth servicing
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 1
    ! Enables WRED on Q1
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 2
! Enables WRED on Q2
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 3
! Enables WRED on Q3
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect min-threshold 1 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q1T1 to 80% and all others to 100%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect max-threshold 1 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q1T1 to 100% and all others to 100%
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect min-threshold 2 80
```

```
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q2T1 to 80% and all others to 100%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect max-threshold 2 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q2T1 to 100% and all others to 100%
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect min-threshold 3 50
60 70 80 90 100 100 100
! Sets Min WRED Threshold for Q3T1 to 50%, Q3T2 to 60%, Q3T3 to 70%
! Q3T4 to 80%, Q3T5 to 90% and all others to 100%
CAT6500-PFC3-IOS(config-if)# wrr-queue random-detect max-threshold 3 60
70 80 90 100 100 100 100
! Sets Max WRED Threshold for Q3T1 to 60%, Q3T2 to 70%, Q3T3 to 80%
! Q3T4 to 90%, Q3T5 to 100% and all others to 100%
CAT6500-PFC3-IOS(config-if)#
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 1 1 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 2 1 0
! Assigns Best Effort to Q2 WRED Threshold 1
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 3 1 4
    ! Assigns Video to Q3 WRED Threshold 1
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 3 2 2
    ! Assigns Net-Mgmt and Transactional Data to Q3 WRED T2
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 3 3 3
! Assigns call signaling and Mission-Critical Data to Q3 WRED T3
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 3 4 6
! Assigns Internetwork-Control (IP Routing) to Q3 WRED T4
CAT6500-PFC3-IOS(config-if)# wrr-queue cos-map 3 5 7
! Assigns Network-Control (Spanning Tree) to Q3 WRED T5
CAT6500-PFC3-IOS(config-if)# priority-queue cos-map 1 5
    ! Assigns VoIP to the PQ (Q4)
CAT6500-PFC3-IOS(config-if)#end
CAT6500-PFC3-IOS#
```

Catalyst 6500 MLS QoS Verification Commands:

- show queueing interface

Catalyst 6500—1P7Q8T Queuing and Dropping

This section includes the following topics:

- Recommendations
- Configuration
- Catalyst MLS QoS Verification Commands

Recommendations

The 1P7Q8T queuing structure adds four additional standard queues to the 1P3Q8T structure and moves the PQ from Q4 to Q8, but otherwise is identical.

Under a 1P7Q8T model, buffer space can be allocated as follows: 5% for the Scavenger/Bulk queue (Q1), 25% for the Best Effort queue (Q2), 10% for the Video queue (Q3), 10% for the Network-Management/Transactional Data queue (Q4), 10% for the Call-Signaling/Mission-Critical Data queue (Q5), 5% for the Internetwork-Control queue (Q6), 5% for the Network Control queue (Q7), and 30% for the PQ (Q8).

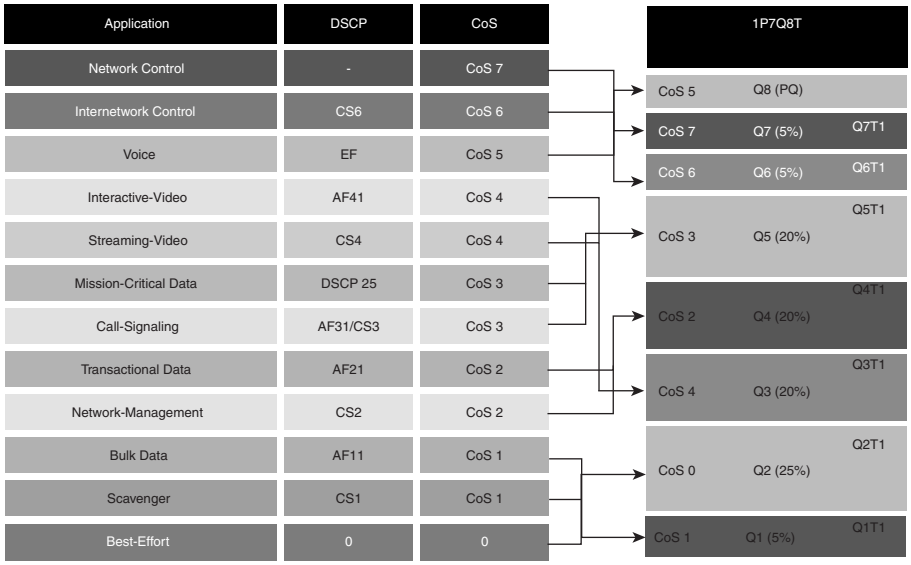
The WRR weights for the standard queues (Q1 through Q7) for dividing the remaining bandwidth after the priority queue has been fully serviced, can be set to 5:25:20:20:20:5:5 respectively for Q1 through Q7.

Since eight queues are available, each CoS value can be assigned to its own exclusive queue. WRED can be enabled on each queue to provide it with congestion avoidance by setting the first WRED threshold of each queue to 80:100. All other WRED thresholds can remain at 100:100.

Therefore, once the queues and thresholds have been defined as above, then CoS 1 (Scavenger/Bulk) can be assigned to Q1T1; CoS 0 (Best Effort) to Q2T1; CoS 4 (Interactive and Streaming Video) can be assigned to Q3T1; CoS 2 (Network Management and Transactional Data) could be assigned to Q4T1; CoS 3 (call signaling and Mission-Critical Data) could be assigned to Q5T1; CoS 6 (Internetwork Control) could be assigned to Q6T1; CoS 7 (Internetwork and Network Control) can be assigned to Q7T1; CoS 5 (VoIP) can be assigned to Q8 (the PQ).

These 1P7Q8T queuing recommendations are illustrated in Figure 2-31.

Figure 2-31 Catalyst 6500 1P7Q8T Queuing Model



Configuration

The Catalyst 6500 (PFC3) CatOS configurations to configure 1P7Q8T queuing recommendations are shown below.

Example 2-77 Catalyst 6500 (PFC3) CatOS—1P7Q8T Queuing Example

```
CAT6500-PFC3-CATOS> (enable) set qos txq-ratio 1p7q8t 5 25 10 10 10 5 5 30
    ! Allocates 5% to Q1, 25% to Q2, 10% to Q3, 10% to Q4,
    ! Allocates 10% to Q5, 5% to Q6, 5% to Q7 and 30% to the PQ (Q8)
CAT6500-PFC3-CATOS> (enable) set qos wrr 1p7q8t 5 25 20 20 20 5 5
! Sets the WRR weights for 5:25:20:20:20:5:5 (Q1 through Q7)
CAT6500-PFC3-CATOS> (enable)
CAT6500-PFC3-CATOS> (enable)
CAT6500-PFC3-CATOS> (enable) set qos wred 1p7q8t tx queue 1 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q1 WRED T1 to 80:100 and all other Q1 WRED Thresholds to 100:100
CAT6500-PFC3-CATOS> (enable) set qos wred 1p7q8t tx queue 2 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q2 WRED T1 to 80:100 and all other Q2 WRED Thresholds to 100:100
CAT6500-PFC3-CATOS> (enable) set qos wred 1p7q8t tx queue 3 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q3 WRED T1 to 80:100 and all other Q3 WRED Thresholds to 100:100
```

```
CAT6500-PFC3-CATOS> (enable) set qos wred 1p7q8t tx queue 4 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q4 WRED T1 to 80:100 and all other Q4 WRED Thresholds to 100:100
CAT6500-PFC3-CATOS> (enable) set qos wred 1p7q8t tx queue 5 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q5 WRED T1 to 80:100 and all other Q5 WRED Thresholds to 100:100
CAT6500-PFC3-CATOS> (enable) set qos wred 1p7q8t tx queue 6 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q6 WRED T1 to 80:100 and all other Q6 WRED Thresholds to 100:100
CAT6500-PFC3-CATOS> (enable) set qos wred 1p7q8t tx queue 7 80:100 100:100
100:100 100:100 100:100 100:100 100:100 100:100
! Sets Q7 WRED T1 to 80:100 and all other Q7 WRED Thresholds to 100:100
CAT6500-PFC3-CATOS> (enable)
CAT6500-PFC3-CATOS> (enable)
CAT6500-PFC3-CATOS> (enable) set qos map 1p7q8t tx 1 1 cos 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
CAT6500-PFC3-CATOS> (enable) set qos map 1p7q8t tx 2 1 cos 0
! Assigns Best Effort to Q2 WRED Threshold 1
CAT6500-PFC3-CATOS> (enable) set qos map 1p7q8t tx 3 1 cos 4
    ! Assigns Video to Q3 WRED Threshold 1
CAT6500-PFC3-CATOS> (enable) set qos map 1p7q8t tx 4 1 cos 2
    ! Assigns Net-Mgmt and Transactional Data to Q4 WRED T1
CAT6500-PFC3-CATOS> (enable) set qos map 1p7q8t tx 5 1 cos 3
! Assigns call signaling and Mission-Critical Data to Q5 WRED T1
CAT6500-PFC3-CATOS> (enable) set qos map 1p7q8t tx 6 1 cos 6
! Assigns Internetwork-Control (IP Routing) to Q6 WRED T1
CAT6500-PFC3-CATOS> (enable) set qos map 1p7q8t tx 7 1 cos 7
! Assigns Network-Control (Spanning Tree) to Q7 WRED T1
CAT6500-PFC3-CATOS> (enable) set qos map 1p7q8t tx 8 cos 5
    ! Assigns VoIP to the PQ (Q4)
CAT6500-PFC3-CATOS> (enable)
```

Catalyst 6500 (PFC3) CatOS QoS Verification Commands:

- show qos info config 1p7q8t tx
- show qos info runtime
- show qos statistics

The Catalyst 6500 (PFC3) IOS configurations to configure 1P7Q8T queuing recommendations are shown below.

Example 2-78 Catalyst 6500 (PFC3) IOS—1P7Q8T Queuing Example

```
CAT6500-PFC3-IOS(config)#interface range TenGigabitEthernet4/1 - 4
CAT6500-PFC3(config-if-range)# wrr-queue queue-limit 5 25 10 10 10 5 5
    ! Allocates 5% to Q1, 25% to Q2, 10% to Q3, 10% to Q4,
    ! Allocates 10% to Q5, 5% to Q6 and 5% to Q7
CAT6500-PFC3(config-if-range)# wrr-queue bandwidth 5 25 20 20 20 5 5
! Sets the WRR weights for 5:25:20:20:20:5:5 (Q1 through Q7)
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 1
    ! Enables WRED on Q1
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 2
! Enables WRED on Q2
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 3
! Enables WRED on Q3
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 4
    ! Enables WRED on Q4
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 5
! Enables WRED on Q5
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 6
```

```
! Enables WRED on Q6
CAT6500-PFC3(config-if-range)# wrr-queue random-detect 7
! Enables WRED on Q7
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)# wrr-queue random-detect min-threshold 1 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q1T1 to 80% and all others to 100%
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect max-threshold 1 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q1T1 to 100% and all others to 100%
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect min-threshold 2 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q2T1 to 80% and all others to 100%
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect max-threshold 2 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q2T1 to 100% and all others to 100%
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect min-threshold 3 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q3T1 to 80% and all others to 100%
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect max-threshold 3 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q3T1 to 100% and all others to 100%
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect min-threshold 4 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q4T1 to 80% and all others to 100%
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect max-threshold 4 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q4T1 to 100% and all others to 100%
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect min-threshold 5 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q5T1 to 80% and all others to 100%
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect max-threshold 5 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q5T1 to 100% and all others to 100%
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect min-threshold 6 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q6T1 to 80% and all others to 100%
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect max-threshold 6 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q6T1 to 100% and all others to 100%
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect min-threshold 7 80
100 100 100 100 100 100 100
! Sets Min WRED Threshold for Q7T1 to 80% and all others to 100%
CAT6500-PFC3(config-if-range)#  wrr-queue random-detect max-threshold 7 100
100 100 100 100 100 100 100
! Sets Max WRED Threshold for Q7T1 to 100% and all others to 100%
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)#
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 1 1 1
! Assigns Scavenger/Bulk to Q1 WRED Threshold 1
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 2 1 0
! Assigns Best Effort to Q2 WRED Threshold 1
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 3 1 4
    ! Assigns Video to Q3 WRED Threshold 1
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 4 1 2
    ! Assigns Net-Mgmt and Transactional Data to Q4 WRED T1
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 5 1 3
```

```
! Assigns call signaling and Mission-Critical Data to Q5 WRED T1
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 6 1 6
! Assigns Internetwork-Control (IP Routing) to Q6 WRED T1
CAT6500-PFC3(config-if-range)# wrr-queue cos-map 7 1 7
! Assigns Network-Control (Spanning Tree) to Q7 WRED T1
CAT6500-PFC3(config-if-range)# priority-queue cos-map 1 5
! Assigns VoIP to the PQ (Q4)
CAT6500-PFC3(config-if-range)#end
CAT6500-PFC3-IOS#
```

Catalyst 6500 MLS QoS Verification Commands:

- show queueing interface

Catalyst 6500—PFC3 Distribution-Layer (IOS) Per-User Microflow Policing

In general, superior defense strategies have multiple lines of defense. In the context of the campus designs that have been considered, there is a main line of defense against DoS/worm attack traffic at the access layer edges. This line of defense can be bolstered at the distribution layer whenever Catalyst 6500 Sup720s (PFC3s) are deployed there. This can be done by leveraging the PFC3 feature of Per-User Microflow Policing.

In the example below, traffic has been assumed to be correctly classified. This may or may not be a valid assumption. If it is suspected to be invalid, then ACLs should be used to identify the flows (instead of DSCP markings). In either case, various flow-types can be filtered as they arrive at the distribution layer to see if they conform to the normal limits that have been set for the enterprise. Each flow is examined by source IP Address and if a source is transmitting out-of-profile, the excess traffic can be dropped or marked-down. In this manner, spurious flows can be contained even in the case that access layer switches do not support granular policing (such as the Catalyst 2950, as discussed earlier in this chapter) or in the case that policing has been mis-configured on an access layer switch.

In this manner, the distribution layer Catalyst 6500 PFC3 can catch any DoS/worm attack flows that may have slipped through the access layer net.

Example 2-79 Catalyst 6500 (PFC3) IOS—Distribution-Layer Per-User Microflow Policing

```
CAT6500-PFC3-IOS(config)#mls qos map policed-dscp normal 0 24 26 34 36 to 8
! Excess traffic marked 0,CS3,AF31,AF41 or AF42 will be remarked to CS1
CAT6500-PFC3-IOS(config)#
CAT6500-PFC3-IOS(config)#class-map match-all VOIP
CAT6500-PFC3-IOS(config-cmap)#  match ip dscp ef
CAT6500-PFC3-IOS(config-cmap)#class-map match-all INTERACTIVE-VIDEO
CAT6500-PFC3-IOS(config-cmap)#  match ip dscp af41 af42
CAT6500-PFC3-IOS(config-cmap)#class-map match-all CALL-SIGNALING
CAT6500-PFC3-IOS(config-cmap)#  match ip dscp cs3 af31
CAT6500-PFC3-IOS(config-cmap)#class-map match-all BEST-EFFORT
CAT6500-PFC3-IOS(config-cmap)#  match ip dscp 0
CAT6500-PFC3-IOS(config-cmap)#
CAT6500-PFC3-IOS(config-cmap)#policy-map PER-USER-POLICING
CAT6500-PFC3-IOS(config-pmap)#  class VOIP
CAT6500-PFC3-I(config-pmap-c)# police flow mask src-only 128000 8000
conform-action transmit exceed-action drop
! No source can send more than 128k worth of DSCP EF traffic
CAT6500-PFC3-I(config-pmap-c)#  class INTERACTIVE-VIDEO
CAT6500-PFC3-I(config-pmap-c)# police flow mask src-only 500000 8000
conform-action transmit exceed-action policed-dscp-transmit
! Excess IP/VC traffic from any source is marked down to CS1
CAT6500-PFC3-I(config-pmap-c)#  class CALL-SIGNALING
CAT6500-PFC3-I(config-pmap-c)# police flow mask src-only 32000 8000
```

```
conform-action transmit exceed-action policed-dscp-transmit
  ! Excess call signaling traffic from any source is marked down to CS1
CAT6500-PFC3-I(config-pmap-c)#  class BEST-EFFORT
CAT6500-PFC3-I(config-pmap-c)# police flow mask src-only 5000000 8000
conform-action transmit exceed-action policed-dscp-transmit
! Excess PC Data traffic from any source is marked down to CS1
CAT6500-PFC3-I(config-pmap-c)#  exit
CAT6500-PFC3-IOS(config-pmap)#exit
CAT6500-PFC3-IOS(config)#
CAT6500-PFC3-IOS(config)#interface range GigabitEthernet4/1 - 4
CAT6500-PFC3(config-if-range)# mls qos trust dscp
CAT6500-PFC3(config-if-range)# service-policy input PER-USER-POLICING
    ! Attaches Per-User Microflow policing policy to Uplinks from Access
CAT6500-PFC3(config-if-range)#end
CAT6500-PFC3-IOS#
```

Catalyst 6500 MLS QoS Verification Commands:

- show mls qos
- show class-map
- show policy-map
- show policy interface

WAN Aggregator/Branch Router Handoff Considerations

A final consideration in campus QoS design is the Campus-to-WAN (or VPN) handoff; in the case of a branch, this equates to the Branch Switch to Branch router handoff.

In either case, a major speed mismatch is impending, as GigabitEthernet/FastEthernet campus networks are connecting to WAN links that may only be a few Megabits (if that).

Granted, the WAN Aggregation Routers and the Remote-Branch Routers have advanced QoS mechanisms to prioritize traffic on their links, but it is critical to keep in mind that Cisco router QoS is performed in IOS *software*, while Catalyst switch QoS is performed in ASIC *hardware*.

Therefore, the optimal distribution of QoS operations would be to have as much QoS actions performed on the Catalyst switches as possible, saving the WAN/Branch router valuable CPU cycles. This is an especially critical consideration when deploying DoS/Worm mitigation designs.

For example, some enterprises have deployed advanced QoS policies on their Branch Switches and Routers, only to have DoS/Worm attacks originate from *within* the Branch. Remember, queuing will not engage on a switch unless its links are congested, and even if it does, should the Branch switch hands off 100 Mbps of (correctly queued) traffic to a Branch router, it will more than likely bring it down.

Thus, the following design principles for the Campus-to-WAN handoff can help mitigate these types of scenarios:

First, resist the urge to automatically use a GigabitEthernet connection to the WAN Aggregation router, even if the router supports GE.

It is extremely unlikely that the WAN Aggregator (WAG) is serving anywhere close to a (combined) WAN-circuit-rate of 1 Gbps. Therefore, use one (or more) FastEthernet connections on the distribution layer Catalyst switch to connect to the WAG, so that the aggregate traffic sent to the WAG is not only *limited* (in 100 Mbps increments), but also (since congestion points are now pulled back into the Catalyst switch, thus forcing queuing to engage on the FE switch port) the traffic will be *correctly queued* within these (100 Mbps-increment) limits.

For example, a WAN Aggregation router may support two DS3 WAN connections (totaling 90 Mbps of WAN circuit-capacity). In this case, the Distribution Layer switch port connecting to the WAG should be FastEthernet. Then, if more than 100 Mbps of traffic attempts to traverse the WAN, the Catalyst switch engages queuing on the switch port and aggressively drops flows according to the defined application hierarchies. Only 100 Mbps of correctly-queued traffic is ever handed off to the WAG.

In the case of a WAG supporting over 100 Mbps of WAN circuits, like the case of a WAG running one or more OC-3 ports (at 155 Mbps each), then multiple FastEthernet connections can be used to connect to the WAG from the Distribution Layer switch to achieve the same net effect.

The point is to bring back, as much as possible, the choke point into Catalyst hardware and engage hardware queuing there, rather than overwhelming the software-based policing and/or queueing policies within the WAG.

Second, if the combined WAN circuit-rate is significantly below 100 Mbps, enable egress shaping on the Catalyst switches (when supported).

If there is no hope of engaging queuing on the Catalyst switch because the combined WAN circuit-rates are far below FastEthernet (the minimum port speed of Catalyst switches), then enable shaping on platforms that support this feature. Such platforms include the Catalyst 2970, 3560, 3750, and 4500.

In this manner, the Catalyst switch can hold back traffic and selectively drop (according to defined policies) from flows that would otherwise flood the WAN/Branch router.

For example, if a Branch router is using two ATM-IMA T1 links (3 Mbps combined throughput) to connect the Branch to the WAN, then the Branch switch could be configured to shape all WAN-destined traffic to 3 Mbps or could be configured to shape on a per-application basis to smaller increments.

Refer to the queuing/dropping sections of these platforms in this chapter and Cisco IOS documentation for additional guidance on enabling shaping.

Finally, if the combined WAN circuit-rate is significantly below 100 Mbps and the Catalyst switch does not support shaping, enable egress policing (when supported).

If the Catalyst switch does not support shaping, then egress policing is the next-best alternative for this scenario.

For example, the Catalyst 3550 does not support shaping, but it does support up to 8 policers on all egress ports. Thus it could still protect its Branch Router from being overwhelmed by policing on egress. Egress policing may be done on an aggregate level or on a per-application-basis.

Again, the objective is to discard, as intelligently as possible, traffic that will inevitably be dropped anyway (by the WAN/Branch router) but, whenever possible, perform the dropping within Catalyst hardware (as opposed to IOS software).

Egress policers are configured in the same manner as ingress policers, but the direction specified in the ***service-policy*** interface-configuration statement will be ***out***, not ***in***.

Note The only Catalyst switch discussed in this chapter that did not support either shaping or egress policing is the Catalyst 2950. Unfortunately there is no way that the Catalyst 2950 can offload QoS from the Branch router. If such functionality is required, then a hardware upgrade would be advisable.

Note For a case study example of Campus QoS design, refer to Figure 12-32 and Examples 12-76 through 12-81 of the Cisco Press book, *End-to-End QoS Network Design* by Tim Szigeti and Christina Hattingh.

Summary

This chapter began with establishing the case for Campus QoS by way of three main QoS design principles:

- The first is that applications should be classified and marked as close to their sources as technically and administratively feasible.
- The second is that unwanted traffic flows should be policed as close to their sources as possible.
- The third is that QoS should always be performed in hardware, rather than software, whenever a choice exists.

Furthermore, it was emphasized that the only way to provide service *guarantees* is to enable queuing at any node that has the potential for congestion, including campus uplinks and downlinks.

A proactive approach to mitigating DoS/worm flooding attacks within campus environments was overviewed. This approach focused on access edge policers that could meter traffic rates received from endpoint devices and when these exceed specified watermarks (at which point they are no longer considered normal flows), these policers could markdown excess traffic to the Scavenger class. These policers would be coupled with queuing policies throughout the enterprise that provisioned for a less-than Best-Effort Scavenger class on all links. In this manner, legitimate traffic bursts would not be affected, but DoS/worm generated traffic would be significantly mitigated.

Common endpoints were overviewed and classified into three main groups: 1) Trusted Endpoints, 2) Untrusted Endpoints, and 3) Conditionally-Trusted Endpoints. Untrusted Endpoints were subdivided into two smaller models: Untrusted PCs and Untrusted Servers; similarly, Conditionally-Trusted Endpoints were subdived into two models: Basic and Advanced.

Following these access edge Model definitions, platform-specific recommendations were given to on how to implement these access edge models on Cisco Catalyst 2950, 2970, 3550, 3560, 3750, 4500, and 6500 series switches. Platform-specific limitations, caveats, or nerd-knobs were highlighted to tailor each model to each platforms unique feature sets. All configurations were presented in config-mode to continually highlight what platform was being discussed. Furthermore, many relevant verification commands were discussed in detail (in context) to illustrate how and when these could be used effectively when deploying QoS within the campus.

Recommendations were also given on how to configure queuing on a per-platform/per-linecard basis. These recommendations included configuring 1P3Q1T queuing on the Catalyst 2950, 1P3Q2T queuing on the Catalyst 3550, configuring 1P3Q3T queuing on the Catalyst 2970/3560/3750, and configuring 1P3Q1T queuing (with DBL) on the Catalyst 4500. For the Catalyst 6500, linecard-specific queuing structures were examined in detail, including CatOS and IOS configurations for configuring 2Q2T, 1P2Q1T, 1P2Q2T, 1P3Q1T, 1P3Q8T, and 1P7Q8T queuing.

Following this, the Catalyst 6500 PFC3's Per-User Microflow Policing feature was discussed in the context of how it could be leveraged to provide a second line of policing defense at the distribution layer.

Finally, Campus-to-WAN/VPN handoff considerations were examined. It was recommended:

- First, resist the urge to automatically use a GigabitEthernet connection to the WAN Aggregation router, even if the router supports GE.
- Second, if the combined WAN circuit-rate is significantly below 100 Mbps, enable egress shaping on the Catalyst switches (when supported).
- Third, if the combined WAN circuit-rate is significantly below 100 Mbps and the Catalyst switch does not support shaping, enable egress policing (when supported).

References

Standards

- RFC 2474 “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers” http://www.ietf.org/rfc/rfc2474
- RFC 2597 “Assured Forwarding PHB Group” http://www.ietf.org/rfc/rfc2597
- RFC 2697 “A Single Rate Three Color Marker” http://www.ietf.org/rfc/rfc2697
- RFC 2698 “A Two Rate Three Color Marker” http://www.ietf.org/rfc/rfc2698
- RFC 3168 “The Addition of Explicit Congestion Notification (ECN) to IP” http://www.ietf.org/rfc/rfc3168
- RFC 3246 “An Expedited Forwarding PHB (Per-Hop Behavior)” http://www.ietf.org/rfc/rfc3246

Books

- Flannagan, Michael and Richard Froom and Kevin Turek. *Cisco Catalyst QoS: Quality of Service in Campus Networks*. Indianapolis: Cisco Press, 2003.
- Szigeti, Tim and Christina Hattingh. *End-to-End QoS Network Design: Quality of Service in LANs, WANs and VPNs*. Indianapolis: Cisco Press, 2004.

Cisco Catalyst Documentation

- Configuring QoS on the Catalyst 2950 (Cisco IOS Software Release 12.1(19)EA1) http://www.cisco.com/univercd/cc/td/doc/product/lan/cat2950/12119ea1/2950scg/swqos.htm
- Configuring QoS on the Catalyst 3550 (Cisco IOS Software Release 12.1(19)EA1) http://www.cisco.com/univercd/cc/td/doc/product/lan/c3550/12119ea1/3550scg/swqos.htm
- Configuring QoS on the Catalyst 2970 (Cisco IOS Software Release 12.2(18)SE) http://www.cisco.com/univercd/cc/td/doc/product/lan/cat2970/12218se/2970scg/swqos.htm
- Configuring QoS on the Catalyst 2970 (Cisco IOS Software Release 12.2(18)SE) http://www.cisco.com/univercd/cc/td/doc/product/lan/cat3750/12218se/3750scg/swqos.htm
- Configuring QoS on the Catalyst 4500 (Cisco IOS Software Release 12.2(18)EW) http://www.cisco.com/univercd/cc/td/doc/product/lan/cat4000/12_2_18/config/qos.htm
- Configuring QoS on the Catalyst 6500 (Cisco CatOS Software Release 8.2) http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/sw_8_2/confg_gd/qos.htm
- Configuring Automatic QoS on the Catalyst 6500 (Cisco CatOS Software Release 8.2) http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/sw_8_2/confg_gd/autoqos.htm
- Configuring QoS on the Catalyst 6500 (Cisco IOS Software Release 12.2(17)SX) http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/122sx/swcg/qos.htm

CHAPTER 3

WAN Aggregator QoS Design

This chapter discusses WAN QoS considerations and designs, including the following:

- Slow-speed (£ 768 kbps) WAN link design
- Medium-speed (768 kbps to T1/E1 speed) WAN link design
- High-speed (> T1/E1 speed) WAN link design

Additionally, these designs are applied to specific Layer 2 WAN media, including the following:

- Leased lines
- Frame Relay
- ATM
- ATM-to-Frame Relay Service Interworking
- ISDN

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.

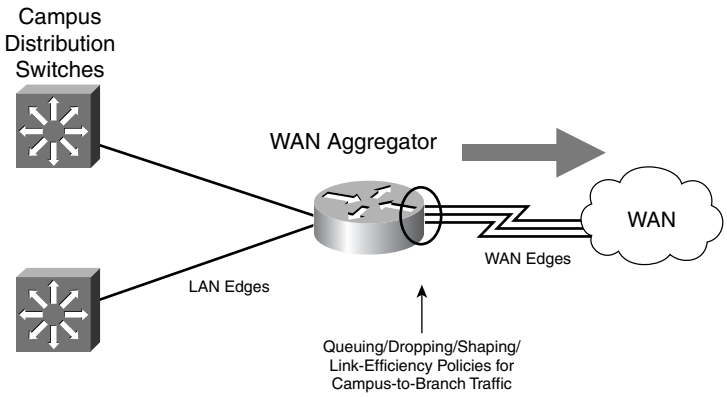
The design principles described in this chapter apply primarily to Layer 2 WANs, such as leased lines, Frame Relay, and ATM (including ATM-to-Frame Relay Service Interworking). However, many service providers use these Layer 2 WAN technologies to access Layer 3 VPN services. Therefore, many of the design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

This chapter provides design guidance for enabling QoS over the WAN. It is important to note that the recommendations in this chapter are not autonomous. They are critically dependent on the recommendations discussed in Chapter 2, "Campus QoS Design."

Where Is QoS Needed over the WAN?

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 chapter 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 Figure 3-1.

Figure 3-1 Where is QoS Needed over the WAN?



WAN Edge QoS Design Considerations

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. Remember, 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. Some of these considerations were introduced in earlier chapters. They are re-emphasized here to underscore their importance to the context of the WAN QoS designs that follow.

Software QoS

Unlike LAN (Catalyst) queuing, which is done in hardware, 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 to keep in mind 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) because this leaves cycles available to respond efficiently to routing updates.

Bandwidth Provisioning for Best-Effort Traffic

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 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. It is recommended that at least 25 percent of a WAN link's bandwidth be reserved for the default Best-Effort class.

Bandwidth Provisioning for Real-Time Traffic

Not only does the Best-Effort class of traffic require special bandwidth-provisioning consideration, but the Real-Time class does as well. The amount of bandwidth assigned to the Real-Time class is variable; however, 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.

The goal of convergence cannot be overemphasized: to enable voice, video, and data to coexist *transparently* on a single network. 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.

Cisco Technical Marketing testing has shown a significant decrease in data application response times when Real-Time traffic exceeds one-third of a link's bandwidth capacity. Cisco IOS Software allows the abstraction (and, thus, configuration) of multiple LLQs. Extensive testing and production-network customer deployments have shown that limiting the sum of all LLQs to 33 percent is a conservative and safe design ratio for merging real-time applications with data applications.

Furthermore, it should be kept in mind that 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 (Š 10 ms per hop) on even medium-speed links
(T1/E1 links) where serialization delays ordinarily would not even be a consideration. (Serialization delays are discussed in more detail in the next section.) Such excessive serialization delays from VoIP LLQ overprovisioning would increase VoIP jitter and, thus, decrease overall call quality.

Note The 33-percent limit for the sum of all LLQs is simply a best-practice design recommendation; it is not a mandate. In some cases, 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 trade-offs involved with overprovisioning LLQ traffic in respect to the negative performance impact on data application response times.

Serialization

Serialization delay refers to the finite amount of time it takes to clock a frame onto the physical media. Within the campus, this time is so infinitesimal that it is completely immaterial. Over the WAN, however, lower link speeds can cause sufficient serialization delay to adversely affect real-time streams, such as Voice or Interactive-Video.

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 (network) delay also is 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 kbps (this is because the serialization delay of a maximum-size Ethernet packet—1500 bytes—takes more than 10 ms to serialize at 768 kbps and below). Naturally, LFI tools need to be enabled on both ends of the link.

When deploying LFI tools, it is recommended that the LFI tools be enabled during a scheduled downtime. Assuming that the network administrator is within the enterprise's 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. Then the administrator can 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 the LLQ escapes fragmentation, it is recommended that Interactive-Video not be deployed 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. Interactive-Video traffic patterns and network requirements are overviewed in Chapter 2, "Campus QoS Design."

IP RTP Header Compression

Compressing IP, UDP, and RTP headers (cRTP) for VoIP calls can result in significant bandwidth gains over WAN links. However, it is important to realize that cRTP is one of the most CPU-intensive features within the Cisco IOS Software QoS toolset. Therefore, it is recommended that cRTP be used primarily on slow-speed (£ 768 kbps) links with a careful eye on CPU levels (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 Example 3-1.

Example 3-1 Displaying the Tx-ring Value with the* show controllers *Command

```
WAG-7206-Left#show controllers Serial 1/0 | include tx_limited
tx_underrun_err=0, tx_soft_underrun_err=0, tx_limited=1(64)
WAG-7206-Left#
```

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 Example 3-2.

Example 3-2 Tuning the Tx-ring

```
WAG-7206-Left(config)#interface Serial 1/0
WAG-7206-Left(config-if)#tx-ring-limit 3
```

The new setting quickly can be verified with the same **show controllers** command, as shown in Example 3-3.

Example 3-3 Verifying Tx-ring Changes

```
WAG-7206-Left#show controllers ser 1/0 | include tx_limited
Tx_underrun_err=0, tx-soft-underru_rr=0, tx-limited=1(3)
WAG-7206_Left#
```

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 kbps links) to approximately 1500 bytes, which is the MTU of Ethernet.

PAK_priority

PAK_priority is the internal Cisco IOS 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/CS6.
- On heavily congested links, it might be necessary to explicitly provision a CBWFQ bandwidth class for routing/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/CS6 because IS-IS uses a CLNS protocol. (It does not use IP, so there are no IPP or DSCP fields to mark.) This is important to keep in mind if explicit bandwidth provisioning is required for IS-IS traffic because it cannot be matched against IPP6/CS6 like most other IGPs. However, NBAR can be used within a class map to match IS-IS traffic (for example, **match protocol clns_is**).
- Although BGPs (both eBGPs and iBGPs) are marked to IPP6/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/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 kbps):
  - 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 a watchful eye on CPU levels).
  - 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 kbps ≤ link speed ≤ T1/E1):
  - VoIP or Interactive-Video can be assigned to the LLQ (usually, there is not enough bandwidth to do both and still keep the 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 (Š 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.

Distributed Platform QoS and Consistent QoS Behavior

It is important to keep in mind that minor differences might exist between QoS configurations on distributed platforms (such as the Cisco 7500 series with VIPs) and those on nondistributed platforms (such as the Cisco 7200 or 1700). The most common difference is the inclusion of the **distributed** keyword after commands such as **ip cef** on distributed platforms. Where more complicated differences exist, they are highlighted explicitly in this chapter.

An important initiative is under way within Cisco to port the QoS code from the Cisco 7500 series routers to the nondistributed router families. This initiative is called Consistent QoS Behavior and has as its objectives simplifying QoS and increasing QoS consistency between platforms. Consistent QoS Behavior code should remove most, if not all, configuration idiosyncrasies between distributed and nondistributed platforms.

WAN Edge Classification and Provisioning Models

One of the most common questions raised when planning a QoS deployment over the WAN is “How many classes of traffic should be provisioned for?” The following considerations should be kept in mind when arriving at an appropriate traffic class model for a given enterprise.

Slow/Medium Link-Speed QoS Class Models

Slow-speed (£ 768 kbps) links have very little bandwidth to carve up, to begin with. When the serialization implications of sending Interactive-Video into the LLQ are taken into consideration, it becomes generally impractical to deploy more than five classes of traffic over slow-speed links.

Medium-speed (£ T1/E1) links do not have serialization restrictions and can accommodate either VoIP or Interactive-Video in their LLQs. However, typically both types of traffic cannot be provisioned at the same time without oversubscribing the LLQ (provisioning more than 33 percent of the traffic for the LLQ). Although this might be possible to configure (the parser will accept the policy and attach it to the interface), the administrator should remember the trade-off of significantly adverse data application response times when LLQs exceed one-third of the link. An alternative approach might be to provision Interactive-Video in a CBWFQ on medium-speed links.

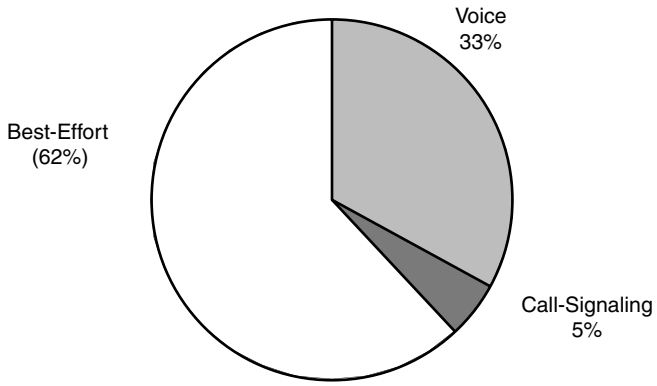
Three-Class (Voice and Data) Model

If the business objective is simply to deploy VoIP over the existing data network, the Voice and Data WAN Edge Model is appropriate. Although it might seem that this is a two-class model, it is actually three: Voice, Call-Signaling, and (generic) data.

Voice is identified by DSCP EF, which is set by default on Cisco IP phones. When identified, VoIP is admitted into the LLQ, which, in this example, is set to the maximum recommended value of 33 percent of the link. Call admission control (CAC) correspondingly should be assigned to this link by dividing the allocated bandwidth by the voice codec (including Layer 2 overhead) to determine how many calls can be permitted simultaneously over this link. Because class-based cRTP is used in this example to compress voice traffic, it also should be factored into the CAC calculation.

Call-Signaling traffic also is marked on the IP phones (to AF31 currently, but it will be migrated to CS3, per the QoS Baseline) and requires a relatively small but dedicated bandwidth guarantee. All other data is fair-queued within class-default. This Three-class WAN Edge Model is illustrated in Figure 3-2 and detailed in Example 3-4.

Figure 3-2 Three-Class WAN Edge Model Migration Strategy Example



Example 3-4 Three-Class WAN Edge Model

```
!
 class-map match-all Voice
  match ip dscp ef            ! IP Phones mark Voice to EF
 class-map match-any Call Signaling
  match ip dscp cs3           ! Future Call-Signaling marking
  match ip dscp af31          ! IP Phones mark Call-Signaling to AF31
!
policy-map WAN-EDGE
  class Voice
   priority percent 33        ! Maximum recommended LLQ value
   compress header ip rtp     ! Optional: Enables Class-Based cRTP
  class Call Signaling
   bandwidth percent 5        ! BW guarantee for Call-Signaling
  class class-default
   fair-queue                 ! All other data gets fair-queuing
!
```

Sometimes administrators explicitly create a class map that functions as the MQC class-default. For instance, an administrator might create a class along the lines of that shown in the following code:

```
class-map match-all BEST-EFFORT
      match any
```

or even:

```
class-map match-all BEST-EFFORT
   match access-group 101
...
access-list 101 permit ip any any
```

These additional configurations are superfluous and inefficient for the router to process. The MQC implicit **class-default** should be used instead.

Another advantage of using the MQC implicit **class-default** is that (currently, before Consistent QoS Behavior code) on nondistributed platforms, class-default is the only class that supports fair queuing within it.

Verification command:

- **show policy**

Verification Command: show policy

The preceding three-class policy, like any other MQC policy, can be verified using the **show policy** command, as shown in Example 3-5.

Example 3-5 Verification of Three-Class WAN Edge Policy

```
RBR-2691-Right#show policy WAN-EDGE
  Policy Map WAN-EDGE
    Class VOICE
      Strict Priority     ! Voice will get LLQ
      Bandwidth 33 (%)    ! LLQ is provisioned to 33%
      compress:
          header ip rtp   ! cRTP is enabled
    Class CALL-SIGNALING
      Bandwidth 5 (%) Max Threshold 64 (packets) ! Call-Signaling gets 5% BW
    Class class-default
      Flow based Fair Queueing                   ! Data will get FQ
      Bandwidth 0 (kbps) Max Threshold 64 (packets)
RBR-2691-Right#
```

The Five-Class WAN Edge Model builds on the previous Three-Class WAN Edge Model and includes a provision for a Critical Data class and a Scavenger class.

The new Critical Data class requires Transactional Data traffic to be marked to DSCP AF21 (or AF22, in the case of dual-rate policers deployed within the campus). Additionally, IGP routing (marked by the routers as CS6) and Network-Management traffic (recommended to be marked to CS2) are protected within this class. In this example, the Critical Data class is provisioned to 36 percent of the link and DSCP-based WRED is enabled on it.

The Scavenger class constrains any traffic marked to DSCP CS1 to 1 percent of the link; this allows class-default to use the remaining 25 percent. However, to constrain Scavenger to 1 percent, an explicit bandwidth guarantee (of 25 percent) must be given to the Best-Effort class. Otherwise, if class-default is not explicitly assigned a minimum bandwidth guarantee, the Scavenger class still can rob it of bandwidth. This is because of the way the CBWFQ algorithm has been coded: If classes protected with a **bandwidth** statement are offered more traffic than their minimum bandwidth guarantee, the algorithm tries to protect such excess traffic at the direct expense of robbing bandwidth from class-default (if class-default is configured with **fair-queue**), *unless* class-default itself has a **bandwidth** statement (providing itself with a minimum bandwidth guarantee). However, assigning a **bandwidth** statement to class-default (on nondistributed platforms) currently precludes the enabling of fair queuing (**fair-queue**) on this class and forces FIFO queuing on class-default (this limitation is to be removed with the release of Consistent QoS Behavior code).

An additional implication of using a **bandwidth** statement on class-default is that even though 25 percent of the link is reserved explicitly for class-default, the parser will not attach the policy to an interface unless the **max-reserved-bandwidth 100** command is entered on the interface before the

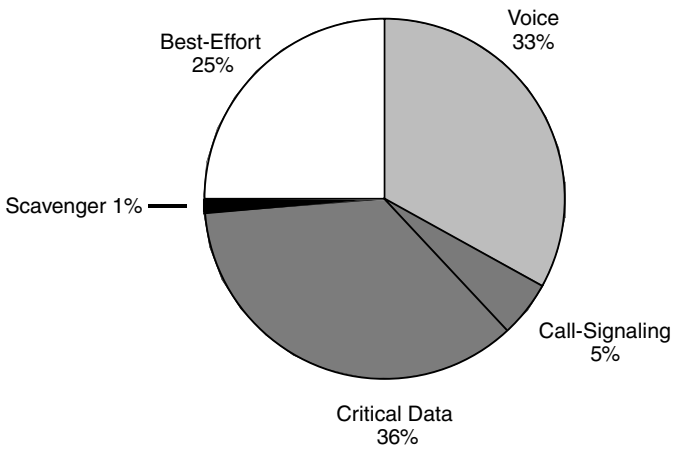
service-policy output statement. This is because the parser adds the sum of the **bandwidth** statements (regardless of whether one of these is applied to the class-default) and, if the total is in excess of 75 percent of the link's bandwidth, rejects the application of the policy to the interface. This is shown in the following code:

```
!
interface Multilink1
 description T1 to Branch#60
 ip address 10.1.112.1 255.255.255.252
 max-reserved-bandwidth 100          ! overrides the default 75% BW limit
 service-policy output WAN-EDGE      ! attaches the MQC policy
 ppp multilink
 ppp multilink group 1
!
```

Furthermore, WRED can be enabled on the Best-Effort class to provide congestion management. Because all traffic assigned to the default class is to be marked to the same DSCP value (of 0), it would be superfluous to enable DSCP-based WRED on such a class; WRED (technically, RED, in this case because all the [IP Precedence] weights are the same) would suffice.

This Five-Class WAN Edge Model is illustrated in Figure 3-3 and detailed in Example 3-6.

Figure 3-3 Five-Class WAN Edge Model Bandwidth Allocation Example



Example 3-6 Five-Class WAN Edge Model

```
!
class-map match-all Voice
  match ip dscp ef                   ! IP Phones mark Voice to EF
class-map match-any Call Signaling
  match ip dscp cs3                  ! Future Call-Signaling marking
  bandwidth percent 1                ! Current Call-Signaling marking
class-map match-any Critical Data
  match ip dscp cs6                  ! Routers mark Routing traffic to CS6
  match ip dscp af21 af22            ! Recommended markings for Transactional-Data
  match ip dscp cs2                  ! Recommended marking for Network Management
class-map match-all Scavenger
  match ip dscp cs1                  ! Scavenger marking
!
 policy-map WAN-EDGE
  class Voice
   priority percent 33               ! Voice gets 33% of LLQ
 class Call Signaling
   bandwidth percent 5               ! BW guarantee for Call-Signaling
```

```
  class Critical Data
   bandwidth percent 36              ! Critical Data class gets 36% BW guarantee
   random-detect dscp-based          ! Enables DSCP-WRED for Critical-Data class
  class Scavenger
   bandwidth percent 1               ! Scavenger class is throttled
  class class-default
   bandwidth percent 25              ! Default class gets a 25% BW guarantee
   random-detect                     ! Enables WRED for class-default
!
```

Verification command:

- **show policy**

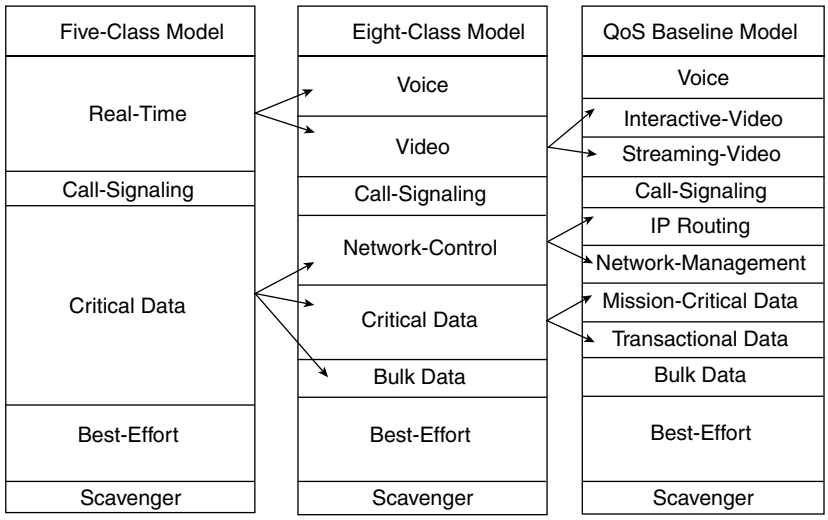
High Link Speed QoS Class Models

High-speed links (such as multiple T1/E1 or above speeds) allow for the provisioning of Voice, Interactive-Video, and multiple classes of data, according to the design rules presented in this chapter (for example, 25 percent for Best Effort class and < 33 percent for all LLQs).

Enabling QoS only optimizes the efficiency of bandwidth utilization; it does not create bandwidth. Therefore, it is important to have adequate bandwidth for all the applications being provisioned. Furthermore, as WAN bandwidth is becoming less expensive, higher-speed links are becoming more popular.

Even if adequate bandwidth exists for up to 11 classes of traffic, as outlined by the QoS Baseline Model, not all enterprises are comfortable with deploying such complex QoS policies at this time. Therefore, it is recommended to start simple, but with room to grow into more complex models. Figure 13-4 illustrates a simple migration strategy showing which classes are good candidates for subdivision into more granular classes as future needs arise.

Figure 3-4 Number of QoS Classes Migration Strategy Example



If the enterprises' QoS requirements exceed that which the Five-Class Model can provision for (such as requiring service guarantees for Interactive-Video and requiring Bulk Data to be controlled during busy periods), they might consider migrating to the Eight-Class Model.

Eight-Class Model

The Eight-Class Model introduces a dual-LLQ design: one for Voice and another for Interactive-Video.

As pointed out in Chapter 5, the LLQ has an implicit policer that allows for time-division multiplexing of the single priority queue. This implicit policer abstracts the fact that there is essentially a single LLQ within the algorithm and, thus, allows for the "provisioning" of multiple LLQs.

Interactive-video (or IP videoconferencing, known also as IP/VC) is recommended to be marked AF41 (which can be marked down to AF42 in the case of dual-rate policing at the campus access edge). It is recommended to overprovision the LLQ by 20 percent of the IP/VC rate. This takes into account IP/UDP/RTP headers as well as Layer 2 overhead.

Additionally, Cisco IOS Software automatically includes a 200-ms burst parameter (defined in bytes) as part of the **priority** command. On dual-T1 links, this has proven sufficient for protecting a single 384-kbps IP/VC stream; on higher-speed links (such as triple T1s), the default burst parameter has shown to be insufficient for protecting multiple IP/VC streams. However, multiple-stream IP/VC quality tested well with the burst set to 30,000 bytes (for example, **priority 920 30000**). Our testing did not arrive at a clean formula for predicting the required size of the burst parameters as IP/VC streams continually were added; however, given the variable packet sizes and rates of these Interactive-Video streams, this is not surprising. The main point is that the default LLQ burst parameter might require tuning as multiple IP/VC streams are added (which likely will be a trial-and-error process).

Optionally, DSCP-based WRED can be enabled on the Interactive-Video class, but testing has shown negligible performance difference in doing so (because, as already has been noted, WRED is more effective on TCP-based flows than UDP-based flows, such as Interactive-Video).

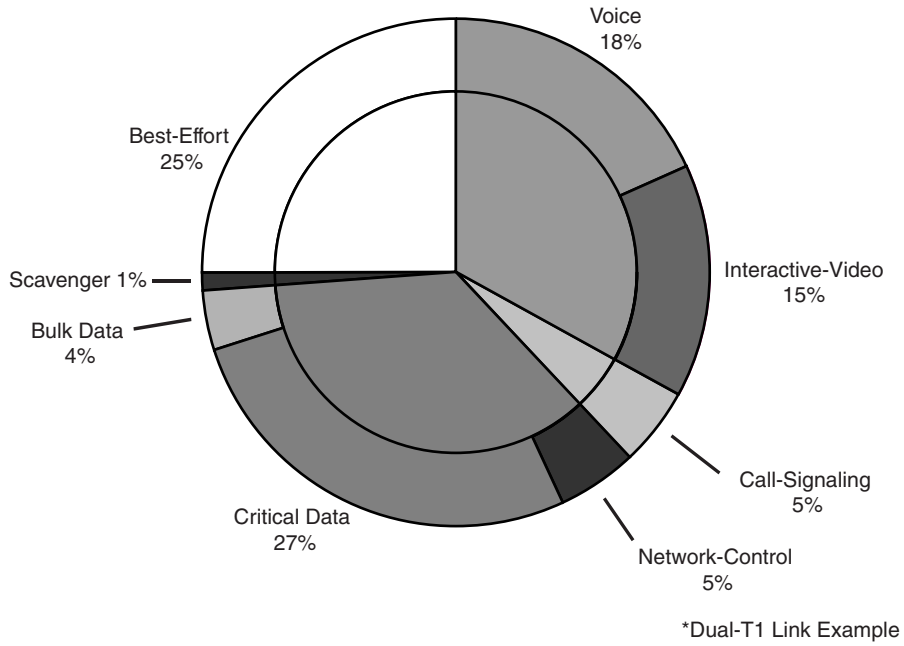
In these designs, WRED is not enabled on classes such as Call-Signaling, IP Routing, or Network-Management because WRED would take effect only if such classes were filling their queues nearly to their limits. Such conditions would indicate a provisioning problem that would better be addressed by increasing the minimum bandwidth allocation for the class than by enabling WRED.

Additionally, the Eight-Class Model subdivides the preferential data class to separate control plane traffic (IP routing and Network-Management applications) from business-critical data traffic. Interior Gateway Protocol (such as RIP, EIGRP, OSPF, and IS-IS) packets are protected through the PAK_priority mechanism within the router. However, EGP protocols, such as BGP, do not get PAK_priority treatment and might need explicit bandwidth guarantees to ensure that peering sessions do not reset during periods of congestion. Additionally, administrators might want to protect network-management access to devices during periods of congestion.

The other class added to this model is for bulk traffic (Bulk Data class), which is also spun away from the Critical Data class. Because TCP continually increases its window sizes, which is especially noticeable in long sessions (such as large file transfers), constraining Bulk Data to its own class alleviates other data classes from being dominated by such large file transfers. Bulk Data is identified by DSCP AF11 (or AF12, in the case of dual-rate policing at the campus access edges). DSCP-based WRED can be enabled on the Bulk Data class (and also on the Critical Data class).

Figure 3-5 shows sample bandwidth allocations of an Eight-Class Model (for a dual-T1 link example). Figure 3-5 also shows how this model can be derived from the Five-Class Model in a manner that maintains respective bandwidth allocations as consistently as possible, which increases the overall end-user transparency of such a migration.

Figure 3-5 Eight-Class WAN Edge Model Bandwidth Allocations Example



Example 3-7 shows the corresponding configuration (over a dual-T1 link) for the Eight-Class Model.

Example 3-7 Eight-Class WAN Edge Model

```
!
class-map match-all Voice
  match ip dscp ef                    ! IP Phones mark Voice to EF
class-map match-all Interactive Video
 match ip dscp af41 af42              ! Recommended markings for IP/VC
class-map match-any Call Signaling
  match ip dscp cs3                   ! Future Call-Signaling marking
  match ip dscp af31                  ! Current Call-Signaling marking
class-map match-any Network Control
  match ip dscp cs6                   ! Routers mark Routing traffic to CS6
  match ip dscp cs2                   ! Recommended marking for Network Management
class-map match-all Critical Data
  match ip dscp af21 af22             ! Recommended markings for Transactional-Data
class-map match-all Bulk Data
  match ip dscp af11 af12             ! Recommended markings for Bulk-Data
class-map match-all Scavenger
  match ip dscp cs1                   ! Scavenger marking
!
 policy-map WAN-EDGE
  class Voice
   priority percent 18        ! Voice gets 552 kbps of LLQ
  class Interactive Video
   priority percent 15        ! 384 kbps IP/VC needs 460 kbps of LLQ
  class Call Signaling
   bandwidth percent 5        ! BW guarantee for Call-Signaling
  class Network Control
   bandwidth percent 5        ! Routing and Network Management get min 5% BW
  class Critical Data
   bandwidth percent 27       ! Critical Data gets min 27% BW
   random-detect dscp-based   ! Enables DSCP-WRED for Critical-Data class
  class Bulk Data
   bandwidth percent 4        ! Bulk Data gets min 4% BW guarantee
```

```
   random-detect dscp-based   ! Enables DSCP-WRED for Bulk-Data class
  class Scavenger
   bandwidth percent 1        ! Scavenger class is throttled
  class class-default
   bandwidth percent 25       ! Fair-queuing is sacrificed for BW guarantee
   random-detect              ! Enables WRED on class-default
!
!
```

Note The Consistent QoS Behavior initiative will enable the configuration of a **bandwidth** statement along with **fair-queue** on any class, including class-default, on all platforms.

Verification command:

- **show policy**

QoS Baseline (11-Class) Model

As mentioned in the overview, the QoS Baseline is a guiding model for addressing the QoS needs of today and the foreseeable future. The QoS Baseline is not a mandate dictating what enterprises must deploy today; instead, this strategic document offers standards-based recommendations for marking and provisioning traffic classes that will allow for greater interoperability and simplified future expansion.

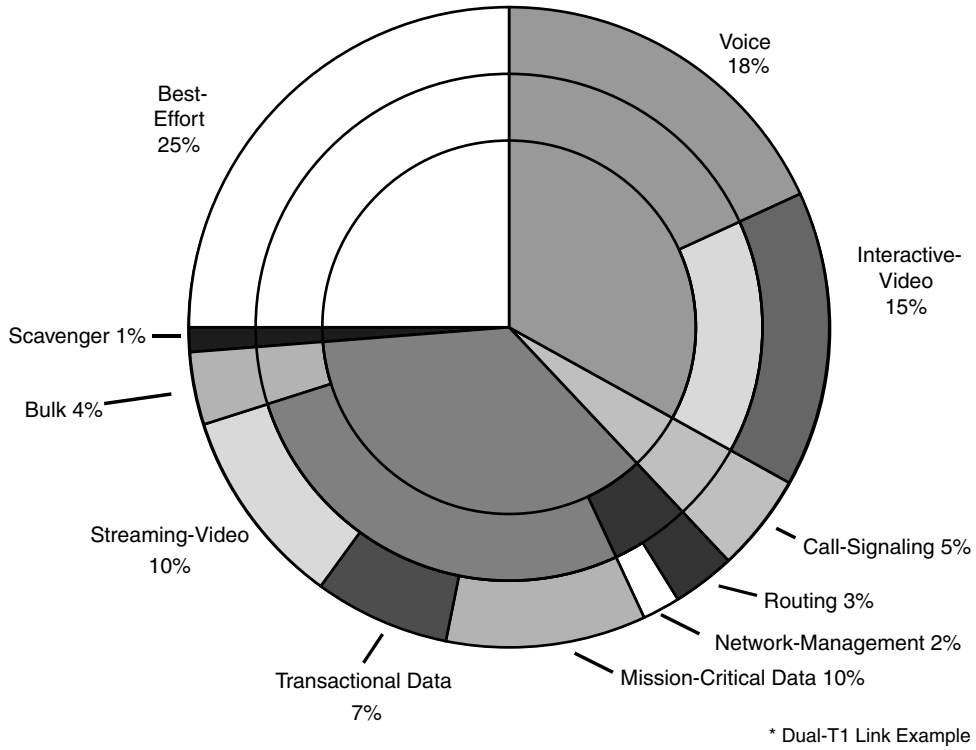
Building on the previous model, the Network-Control class is subdivided into the IP Routing and Network-Management classes.

The Critical Data class also is subdivided further into the Mission-Critical Data and Transactional Data classes. Although DSCP-based WRED is enabled on the Transactional Data class, because packets for this class can be marked AF21 (or AF22, as in the case of dual-rate policers being deployed in the campus), it would be superfluous to enable DSCP-based WRED on the Mission-Critical Data class (WRED will suffice because all Mission-Critical Data class packets are marked to the same value: DSCP 25).

Finally, a new class is provisioned for Streaming-Video. Testing has shown that there is a negligible difference in enabling WRED on this UDP-based traffic class, so, although it remains an option, WRED is not enabled in these design examples.

Figure 3-6 shows a sample WAN edge bandwidth allocation for a QoS Baseline Model (over a dual-T1 link) and also shows how this model can be derived from the Five- and Seven-Class Models in a manner that maintains respective bandwidth allocations as consistently as possible. This increases the overall end-user transparency of such a migration.

Figure 3-6 QoS Baseline WAN Edge Model Bandwidth Allocations Example



Example 3-8 shows the corresponding configuration for an 11-Class QoS Baseline WAN Edge Model (over a dual-T1 link).

Example 3-8 QoS Baseline WAN Edge Model

```
!
class-map match-all Voice
  match ip dscp ef                    ! IP Phones mark Voice to EF
class-map match-all Interactive Video
  match ip dscp af41 af42             ! Recommended markings for IP/VC
class-map match-any Call Signaling
  match ip dscp cs3                   ! Future Call-Signaling marking
  match ip dscp af31                  ! Current Call-Signaling marking
class-map match-all Routing
  match ip dscp cs6                   ! Routers mark Routing traffic to CS6
class-map match-all Net Mgmt
  match ip dscp cs2                   ! Recommended marking for Network Management
class-map match-all Mission-Critical Data
  match ip dscp 25                    ! Interim marking for Mission-Critical Data
class-map match-all Transactional Data
  match ip dscp af21 af22             ! Recommended markings for Transactional-Data
class-map match-all Bulk Data
  match ip dscp af11 af12             ! Recommended markings for Bulk-Data
class-map match-all Streaming Video
  match ip dscp cs4                   ! Recommended marking for Streaming-Video
class-map match-all Scavenger
  match ip dscp cs1                   ! Recommended marking for Scavenger traffic
!
 policy-map WAN-EDGE
  class Voice
   priority percent 18        ! Voice gets 552 kbps of LLQ
  class Interactive Video
```

```
   priority percent 15        ! 384 kbps IP/VC needs 460 kbps of LLQ
  class Call Signaling
   bandwidth percent 5        ! BW guarantee for Call-Signaling
  class Routing
   bandwidth percent 3        ! Routing class gets explicit BW guarantee
  class Net Mgmt
   bandwidth percent 2        ! Net-Mgmt class gets explicit BW guarantee
  class Mission-Critical Data
   bandwidth percent 10       ! Mission-Critical class gets 10% BW guarantee
   random-detect              ! Enables WRED for Mission-Critical Data class
  class Transactional Data
   bandwidth percent 7        ! Transactional-Data class gets 7% BW guarantee
   random-detect dscp-based   ! Enables DSCP-WRED for Transactional-Data class
  class Bulk Data
   bandwidth percent 4        ! Bulk Data remains at 4% BW guarantee
   random-detect dscp-based   ! Enables DSCP-WRED for Bulk-Data class
  class Streaming Video
   bandwidth percent 10       ! Streaming-Video class gets 10% BW guarantee
  class Scavenger
   bandwidth percent 1        ! Scavenger class is throttled
  class class-default
   bandwidth percent 25       ! Class-Default gets 25% min BW guarantee
   random-detect              ! Enables WRED on class-default
!
```

Verification command:

- **show policy**

Again, a **bandwidth** statement is used on class-default (currently), precluding the use of **fair-queue** on the class for all nondistributed platforms. Also, a **max-reserved-bandwidth 100** statement must be applied to the interface before the **service-policy output** statement.

Distributed-Platform/Consistent QoS Behavior—QoS Baseline Model

One of the current advantages of the Cisco 7500 (distributed platform) QoS code is that it can support **bandwidth** commands in conjunction with **fair-queue** on any given class, including class-default. This functionality will become available to nondistributed platforms with the release of Consistent QoS Behavior code. (As of this writing, this initiative does not have a fixed target delivery date.) When **fair-queue** is enabled on the main data classes, the resulting configuration becomes as shown in Example 3-9.

Example 3-9 Distributed-Platform/Consistent QoS Behavior—QoS Baseline WAN Edge Model

```
!
ip cef distributed       ! 'distributed' keyword required on 7500 for ip cef
!
class-map match-all Voice
  match ip dscp ef                  ! IP Phones mark Voice to EF
class-map match-all Interactive Video
  match ip dscp af41 af42           ! Recommended markings for IP/VC
class-map match-any Call Signaling
  match ip dscp cs3                 ! Future Call-Signaling marking
  match ip dscp af31                ! Current Call-Signaling marking
class-map match-all Routing
  match ip dscp cs6                 ! Routers mark Routing traffic to CS6
class-map match-all Net Mgmt
  match ip dscp cs2                 ! Recommended marking for Network Management
class-map match-all Mission-Critical Data
  match ip dscp 25                  ! Interim marking for Mission-Critical Data
class-map match-all Transactional Data
  match ip dscp af21 af22           ! Recommended markings for Transactional-Data
```

```
class-map match-all Bulk Data
  match ip dscp af11 af12            ! Recommended markings for Bulk-Data
class-map match-all Streaming Video
  match ip dscp cs4                  ! Recommended marking for Streaming-Video
class-map match-all Scavenger
  match ip dscp cs1                  ! Recommended marking for Scavenger traffic
!
policy-map WAN-EDGE
  class Voice
   priority percent 18        ! Voice gets 552 kbps of LLQ
  class Interactive Video
   priority percent 15        ! 384 kbps IP/VC needs 460 kbps of LLQ
  class Call Signaling
   bandwidth percent 5        ! Bandwidth guarantee for Call-Signaling
  class Routing
   bandwidth percent 3        ! Bandwidth guarantee for Routing
  class Net Mgmt
   bandwidth percent 2        ! Bandwidth guarantee for Network Management
  class Mission-Critical Data
   bandwidth percent 10       ! Mission-Critical data gets min 10% BW guarantee
   fair-queue                 ! Applies FQ to Mission-Critical Data class
   random-detect              ! Enables WRED on Mission-Critical Data class
  class Transactional Data
   bandwidth percent 7        ! Transactional Data gets min 7% BW guarantee
   fair-queue                 ! Applies FQ to Transactional Data class
   random-detect dscp-based   ! Enables DSCP-WRED on Transactional Data class
  class Bulk Data
   bandwidth percent 4        ! Bulk Data gets min 4% BW guarantee
   fair-queue                 ! Applies FQ to Bulk Data class
   random-detect dscp-based   ! Enables DSCP-WRED on Bulk Data class
  class Streaming Video
   bandwidth percent 10       ! Streaming-Video gets min 10% BW guarantee
  class Scavenger
   bandwidth percent 1        ! Scavenger class is throttled
  class class-default
   bandwidth percent 25       ! Class-Default gets min 25% BW guarantee
   fair-queue                 ! Applies FQ to Class-Default
   random-detect              ! Enables WRED on Class-Default
!
```

WAN Edge Link-Specific QoS Design

The most popular WAN media in use today are leased lines, Frame Relay, and ATM (including ATM-to-Frame Relay Service Interworking). Each of these media can be deployed in three broad categories of link speeds: slow speed (£ 768 kbps), medium speed (£ T1/E1), and high speed (multiple T1/E1 or greater). The following sections detail specific designs for each medium at each speed category. Additionally, ISDN QoS design is discussed in the context of a backup WAN link.

Leased Lines

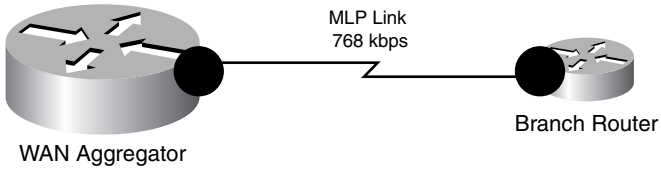
Leased lines, or point-to-point links, can be configured with HDLC, PPP, or MLP encapsulation. MLP offers the network administrator the most flexibility and deployment options. For example, MLP is the only leased-line protocol that supports LFI on slow-speed links (through MLP LFI). Additionally, as bandwidth requirements grow over time, MLP requires the fewest modifications to accommodate the addition of multiple T1/E1 lines to a WAN link bundle. Furthermore, MLP supports all of the security options of PPP (such as CHAP authentication).

Slow-Speed (£768 kbps) Leased Lines

Recommendation: Use MLP LFI and cRTP.

For slow-speed leased lines (as illustrated in Figure 3-7), LFI is required to minimize serialization delay. MLP, therefore, is the only encapsulation option on slow-speed leased lines because MLP LFI is the only mechanism available for fragmentation and interleaving on such links. Optionally, cRTP can be enabled either as part of the MQC policy map
(as shown in Example 3-10) or under the multilink interface (using the **ip rtp header-compression** command). Ensure that MLP LFI and cRTP, if enabled, are configured on both ends of the point-to-point link, as shown in Example 3-14.

Figure 3-7 Slow-Speed Leased Lines



Example 3-10 Slow-Speed (£768 kbps) Leased-Line QoS Design Example

```
!
policy-map WAN-EDGE
  class Voice
   priority percent 33         ! Maximum recommended LLQ value
   compress header ip rtp      ! Enables Class-Based cRTP
  class Call Signaling
   bandwidth percent 5         ! BW guarantee for Call-Signaling
…                              ! A 3 to 5 Class Model can be used
!
interface Multilink1
 description 768 kbps Leased-Line to RBR-3745-Left
 ip address 10.1.112.1 255.255.255.252
 service-policy output WAN-EDGE  ! Attaches the MQC policy to Mu1
 ppp multilink
 ppp multilink fragment delay 10  ! Limits serialization delay to 10 ms
 ppp multilink interleave         ! Enables interleaving of Voice with Data
 ppp multilink group 1
!
…
!
interface Serial1/0
 bandwidth 786
 no ip address
 encapsulation ppp
 ppp multilink
 ppp multilink group 1          ! Includes interface Ser1/0 into Mu1 group
!
```

Verification commands:

- **show policy**
- **show interface**
- **show policy interface**
- **show ppp multilink**

Verification Command: show interface

The **show interface** command indicates whether drops are occurring on an interface (an indication of congestion). Additionally, on a multilink interface with LFI enabled, the command displays interleaving statistics, as shown in Example 3-11.

Example 3-11 **show interface** ***Verification of MLP LFI on a Slow-Speed Leased Line***

```
WAG-7206-Left#show interface multilink 1
Multilink1 is up, line protocol is up
  Hardware is multilink group interface
  Description: 768 kbps Leased-Line to RBR-3745-Left
  Internet address is 10.1.112.1/30
  MTU 1500 bytes, BW 768 Kbit, DLY 100000 usec,
     reliability 255/255, txload 233/255, rxload 1/255
  Encapsulation PPP, LCP Open, multilink Open
  Open: CDPCP, IPCP, loopback not set
  DTR is pulsed for 2 seconds on reset
  Last input 00:00:01, output never, output hang never
  Last clearing of "show interface" counters 00:16:15
  Input queue: 0/75/0/0 (size/max/drops/flushes);
  Total output drops: 49127
  Queueing strategy: weighted fair
  Output queue: 54/1000/64/49127/185507
  (size/max total/threshold/drops/interleaves)
```

In Example Example 3-11, 49,127 drops have occurred on the multilink interface (because of congestion), and LFI has engaged with 185,507 interleaves of voice with data.

Verification Command: **show policy interface** (Three-Class Policy)

The **show policy interface** command is probably the most useful **show** command for MQC-based QoS policies. It displays a wide array of dynamic statistics, including the number of matches on a class map as a whole, the number of matches against each discrete **match** statement within a class map, the number of queued or dropped packets (either tail dropped or WRED dropped), and many other relevant QoS statistics. Example 3-12 shows example output of the **show policy interface** command.

Example 3-12 **show policy interface** ***Verification of a Three-Class Policy on a Slow-Speed Leased Line***

```
WAG-7206-Left#show policy interface multilink 1
 Multilink1
  Service-policy output: WAN-EDGE
    Class-map: Voice (match-all)
      68392 packets, 4377088 bytes
      30 second offered rate 102000 bps, drop rate 0 bps
      Match: ip dscp ef
      Queueing
        Strict Priority

        Output Queue: Conversation 264
        Bandwidth 33 (%)
        Bandwidth 253 (kbps) Burst 6325 (Bytes)
        (pkts matched/bytes matched) 68392/2043848
        (total drops/bytes drops) 0/0
      compress:
          header ip rtp
          UDP/RTP compression:
          Sent: 68392 total, 68388 compressed,
                2333240 bytes saved, 1770280 bytes sent
                2.31 efficiency improvement factor
                99% hit ratio, five minute miss rate 0 misses/sec,0 max
```

```
              rate 41000 bps
    Class-map: Call Signaling (match-any)
      251 packets, 142056 bytes
      30 second offered rate 3000 bps, drop rate 0 bps
      Match: ip dscp cs3
        0 packets, 0 bytes
        30 second rate 0 bps
      Match: ip dscp af31
        251 packets, 142056 bytes
        30 second rate 3000 bps
      Queueing
        Output Queue: Conversation 265
        Bandwidth 5 (%)
        Bandwidth 38 (kbps) Max Threshold 64 (packets)
        (pkts matched/bytes matched) 255/144280
        (depth/total drops/no-buffer drops) 0/0/0
    Class-map: class-default (match-any)
      51674 packets, 28787480 bytes
      30 second offered rate 669000 bps, drop rate 16000 bps
      Match: any
      Queueing
        Flow Based Fair Queueing
        Maximum Number of Hashed Queues 256
        (total queued/total drops/no-buffer drops) 36/458/0
WAG-7206-Left#
```

In Example 3-12, the Voice class map and Call-Signaling class map are receiving matches on their classification criteria (DSCP EF and DSCP CS3/AF31, respectively). However, because Cisco IP Telephony products currently mark Call-Signaling traffic to DSCP AF31, Call-Signaling traffic is matching only on DSCP AF31 in this example.

The last line of every class map output is important because this line indicates whether any drops are occurring on this traffic class. In this example, there are no drops in the Voice or Call-Signaling classes, which is the desired behavior. A few drops are occurring in class-default, but this is expected when the interface is congested (which is the trigger to engage queuing).

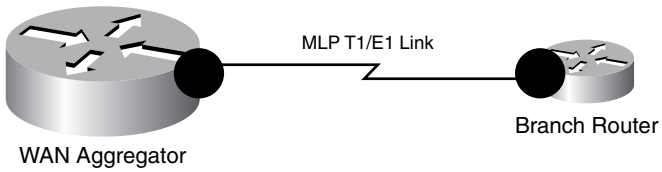
Also of note, and specific to this particular configuration, are the cRTP statistics included under the Voice class map. These cRTP statistics are displayed because class-based cRTP was enabled in this example (instead of enabling cRTP on the interface). Remember, cRTP must be enabled on both ends of the links for compression to occur; otherwise, these counters will never increment.

Medium-Speed (£ T1/E1) Leased Lines

Recommendation: MLP LFI is not required; cRTP is optional.

Medium-speed leased lines (as shown in Figure 3-8) can use HDLC, PPP, or MLP encapsulation. An advantage of using MLP encapsulation is that future growth (to multiple T1/E1 links) will be easier to manage. Also, MLP includes all the security options of PPP (such as CHAP).

Figure 3-8 Medium-Speed Leased Lines



However, MLP LFI is not required at these speeds, and cRTP is optional. Example 3-13 shows an example configuration for medium-speed leased lines.

Example 3-13 Medium-Speed Leased-Line QoS Design Example

```
!
interface Multilink1
 description T1 Leased-Line to RBR-3745-Left
 ip address 10.1.112.1 255.255.255.252
 service-policy output WAN-EDGE     ! Attaches the MQC policy to Mu1
 ppp multilink
 ppp multilink group 1       ! Identifies Mu1 as logical Int for Mu1 group
!
…
!
interface Serial1/0
 bandwidth 1536
 no ip address
 encapsulation ppp
 load-interval 30
 ppp multilink
 ppp multilink group 1        ! Includes interface Ser1/0 into Mu1 group
!
```

Verification commands:

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

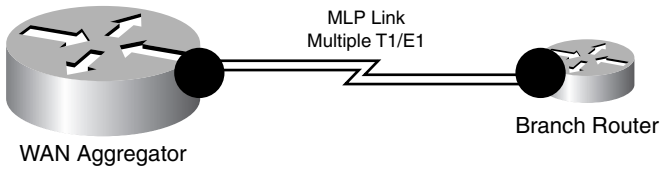
High-Speed (Multiple T1/E1 or Greater) Leased Lines

Recommendation: Use MLP bundling, but keep an eye on CPU levels. When enterprises have multiple T1/E1-speed leased lines to individual branches, three options exist for load sharing:

- IP CEF per-destination load balancing
- IP CEF per-packet load balancing
- Multilink PPP bundles

Cisco Technical Marketing testing has shown that IP CEF per-destination load balancing does not meet the SLAs required for Voice and Interactive-Video over multiple T1/E1 links, as shown in Figure 3-9.

Figure 3-9 High-Speed Leased Lines



On the other hand, IP-CEF per-packet load balancing did meet the required SLAs, but not quite as well as MLP bundling.

MLP bundling attained the best overall SLA values for delay and jitter, but it required more CPU resources than IP CEF per-packet load balancing. If CPU levels are kept under the recommended 75 percent, it is recommended to use MLP bundling for multiple T1/E1 links.

Also, if policy maps that require bandwidth statements on class-default are being attached to the multilink interface, the **max-reserved-bandwidth 100** command is required on the interface before the **service-policy output** statement can be applied, as shown in Example 3-14.

Example 3-14 High-Speed (≠ Multiple T1/E1) Leased Line QoS Design Example

```
!
interface Multilink1
 description Dual-T1 to RBR-3745-Left
 ip address 10.1.112.1 255.255.255.252
 max-reserved-bandwidth 100        ! Overrides the default 75% BW limit
 service-policy output WAN-EDGE    ! Attaches the MQC policy to Mu1
 ppp multilink
 ppp multilink group 1       ! Identifies Mu1 as logical int for Mu1 group
!
…
!
interface Serial1/0
 bandwidth 1536        ! defined on physical interface only
 no ip address
 encapsulation ppp
 ppp multilink
 ppp multilink group 1        ! includes interface Ser1/0 into Mu1 group
!
interface Serial1/1
 bandwidth 1536        ! defined on physical interface only
 no ip address
 encapsulation ppp
 ppp multilink
 ppp multilink group 1        ! includes interface Ser1/1 into Mu1 group
!
```

Note Interface **bandwidth** commands (not to be confused with policy map CBWFQ **bandwidth** commands) should be defined only on the physical interfaces, not on multilink interfaces. This way, if any physical interfaces go down, the Cisco IOS Software will reflect the change in the multilink interface's bandwidth for routing and QoS purposes. This change can be verified by the **show interface** command. However, if a bandwidth statement is configured under the multilink interface, the bandwidth value for the interface will be static even if an underlying physical interface is lost.

Verification commands:

- **show policy**
- **show interface**
- **show policy interface**
- **show ppp multilink**

Verification Command: show policy interface (QoS Baseline Policy)

A more complex example of the **show policy interface** command is given in Example 3-15, where a QoS Baseline WAN edge policy is being applied to a dual-T1 (high-speed) leased line.

Example 3-15 show policy interface Verification of a QoS Baseline Policy on a High-Speed Leased Line

```
WAG-7206-Left#show policy interface multilink 1
```

```
Multilink1
 Service-policy output: WAN-EDGE
   Class-map: Voice (match-all)
     444842 packets, 28467338 bytes
     30 second offered rate 434000 bps, drop rate 0 bps
     Match: ip dscp ef
     Queueing
       Strict Priority
       Output Queue: Conversation 264
       Bandwidth 18 (%)
       Bandwidth 552 (kbps) Burst 13800 (Bytes)
       (pkts matched/bytes matched) 444842/28467338
       (total drops/bytes drops) 0/0
   Class-map: Interactive Video (match-all)
     32685 packets, 25977946 bytes
     30 second offered rate 405000 bps, drop rate 0 bps
     Match: ip dscp af41
     Queueing
       Strict Priority
       Output Queue: Conversation 264
       Bandwidth 15 (%)
       Bandwidth 460 (kbps) Burst 11500 (Bytes)
       (pkts matched/bytes matched) 32843/26097186
       (total drops/bytes drops) 0/0
   Class-map: Call Signaling (match-any)
     1020 packets, 537876 bytes
     30 second offered rate 7000 bps, drop rate 0 bps
     Match: ip dscp cs3
       0 packets, 0 bytes
       30 second rate 0 bps
     Match: ip dscp af31
       1020 packets, 537876 bytes
       30 second rate 7000 bps
     Queueing
       Output Queue: Conversation 265
       Bandwidth 5 (%)
       Bandwidth 153 (kbps) Max Threshold 64 (packets)
       (pkts matched/bytes matched) 1022/538988
       (depth/total drops/no-buffer drops) 0/0/0
   Class-map: Routing (match-all)
     1682 packets, 112056 bytes
     30 second offered rate 0 bps, drop rate 0 bps
     Match: ip dscp cs6
     Queueing
       Output Queue: Conversation 266
       Bandwidth 3 (%)
       Bandwidth 92 (kbps) Max Threshold 64 (packets)
       (pkts matched/bytes matched) 1430/95844
       (depth/total drops/no-buffer drops) 0/0/0
   Class-map: Net Mgmt (match-all)
     32062 packets, 2495021 bytes
     30 second offered rate 41000 bps, drop rate 0 bps
     Match: ip dscp cs2
     Queueing
       Output Queue: Conversation 267
       Bandwidth 2 (%)
       Bandwidth 61 (kbps) Max Threshold 64 (packets)
       (pkts matched/bytes matched) 32256/2510284
       (depth/total drops/no-buffer drops) 0/0/0
   Class-map: Mission-Critical Data (match-all)
     56600 packets, 40712013 bytes
     30 second offered rate 590000 bps, drop rate 0 bps
     Match: ip dscp 25
     Queueing
```

```
        Output Queue: Conversation 268
        Bandwidth 12 (%)
        Bandwidth 368 (kbps)
        (pkts matched/bytes matched) 57178/41112815
        (depth/total drops/no-buffer drops) 10/0/0
         exponential weight: 9
         mean queue depth: 10
class   Transmitted    Random drop    Tail drop   Minimum Maximum  Mark
        pkts/bytes     pkts/bytes     pkts/bytes  thresh  thresh   prob
0          0/0            0/0            0/0        20      40     1/10
1          0/0            0/0            0/0        22      40     1/10
2          0/0            0/0            0/0        24      40     1/10
3      57178/41112815     0/0            0/0        26      40     1/10
4          0/0            0/0            0/0        28      40     1/10
5          0/0            0/0            0/0        30      40     1/10
6          0/0            0/0            0/0        32      40     1/10
7          0/0            0/0            0/0        34      40     1/10
rsvp       0/0            0/0            0/0        36      40     1/10
    Class-map: Transactional Data (match-all)
      31352 packets, 31591979 bytes
      30 second offered rate 435000 bps, drop rate 10000 bps
      Match: ip dscp af21
      Queueing
        Output Queue: Conversation 269
        Bandwidth 8 (%)
        Bandwidth 245 (kbps)
        (pkts matched/bytes matched) 31741/32008133
        (depth/total drops/no-buffer drops) 29/954/0
         exponential weight: 9
         mean queue depth: 26

class   Transmitted    Random drop    Tail drop   Minimum Maximum  Mark
        pkts/bytes     pkts/bytes     pkts/bytes  thresh  thresh   prob
0          0/0            0/0            0/0        20      40     1/10
1          0/0            0/0            0/0        22      40     1/10
2      30787/31019741   954/988392       0/0        24      40     1/10
3          0/0            0/0            0/0        26      40     1/10
4          0/0            0/0            0/0        28      40     1/10
5          0/0            0/0            0/0        30      40     1/10
6          0/0            0/0            0/0        32      40     1/10
7          0/0            0/0            0/0        34      40     1/10
rsvp       0/0            0/0            0/0        36      40     1/10
    Class-map: Streaming Video (match-all)
      23227 packets, 19293728 bytes
      30 second offered rate 291000 bps, drop rate 0 bps
      Match: ip dscp cs4
      Queueing
        Output Queue: Conversation 271
        Bandwidth 10 (%)
        Bandwidth 307 (kbps) Max Threshold 64 (packets)
        (pkts matched/bytes matched) 23683/19672892
        (depth/total drops/no-buffer drops) 2/0/0

  Class-map: Scavenger (match-all)
      285075 packets, 129433625 bytes
      30 second offered rate 2102000 bps, drop rate 2050000 bps
      Match: ip dscp cs1
      Queueing
        Output Queue: Conversation 272
        Bandwidth 1 (%)
        Bandwidth 30 (kbps) Max Threshold 64 (packets)
        (pkts matched/bytes matched) 291885/132532775
        (depth/total drops/no-buffer drops) 64/283050/0
    Class-map: class-default (match-any)
```

```
      40323 packets, 35024924 bytes
      30 second offered rate 590000 bps, drop rate 0 bps
      Match: any
      Queueing
        Output Queue: Conversation 273
        Bandwidth 25 (%)
        Bandwidth 768 (kbps)
        (pkts matched/bytes matched) 41229/35918160
        (depth/total drops/no-buffer drops) 12/268/0
         exponential weight: 9
         mean queue depth: 4
class   Transmitted     Random drop     Tail drop    Minimum Maximum  Mark
        pkts/bytes      pkts/bytes      pkts/bytes   thresh  thresh   prob
0      40961/35700528   268/217632        0/0         20      40      1/10
1          0/0              0/0             0/0         22      40      1/10
2          0/0              0/0             0/0         24      40      1/10
3          0/0              0/0             0/0         26      40      1/10
4          0/0              0/0             0/0         28      40      1/10
5          0/0              0/0             0/0         30      40      1/10
6          0/0              0/0             0/0         32      40      1/10
7          0/0              0/0             0/0         34      40      1/10
rsvp       0/0              0/0             0/0         36      40      1/10
```

Important items to note for a given class are the **pkts matched** statistics (which verify that classification has been configured correctly and that the packets have been assigned to the proper queue) and the **total drops** statistics (which indicate whether adequate bandwidth has been assigned to the class).

Extremely few drops, if any, are desired in the Voice, Interactive-Video, Call-Signaling, and Routing classes.

Note The Routing class is a special case because of the statistics that it displays.

On nondistributed platforms, the classification counter (the first line under the class map) shows any IGP traffic matched by the Routing class (identified by DSCP CS6). But remember that IGP protocols queue separately (because these are handled by the PAK_priority mechanism) and, therefore, do not register queuing statistics within the MQC counters for the Routing class. EGP protocols (such as BGP), on the other hand, do register queuing/dropping statistics within such an MQC class.

The situation is different on distributed platforms, where all routing packets (IGP or EGP) are matched and queued within a provisioned Routing class (complete with queuing/dropping statistics through the **show policy interface** verification command).

Few drops are expected in the Mission-Critical Data class. WRED (essentially RED because all packets are marked to the same IPP/DSCP value) is enabled to avoid congestion on this class. Some drops are expected for the Transactional Data class, yet, in this particular example, WRED is minimizing tail drops for this class.

It is normal for the Bulk Data class to show drops (both WRED and tail). This is because the Bulk Data class is being constrained from dominating bandwidth by its large and sustained TCP sessions. The Scavenger class should show very aggressive dropping during periods of congestion. Finally, it is normal for drops to appear in the default class.

Verification Command: **show ppp multilink**

The **show ppp multilink** command is useful to verify that multiple physical links are correctly associated and included in the MLP bundle, as shown in Example 3-16. Also, the load (which might not quite hit 255/255) indicates congestion on the link.

Example 3-16 **show ppp multilink** ***Verification of a High-Speed Leased Line***

```
WAG-7206-Left#show ppp multilink
Multilink1, bundle name is RBR-3745-Left
  Bundle up for 00:28:33, 254/255 load
  Receive buffer limit 24384 bytes, frag timeout 1000 ms
    0/0 fragments/bytes in reassembly list
    0 lost fragments, 2 reordered
    0/0 discarded fragments/bytes, 0 lost received
    0xE8F received sequence, 0x9A554 sent sequence
  Member links: 2 active, 0 inactive (max not set, min not set)
    Se1/0, since 00:28:35, 1920 weight, 1496 frag size
    Se1/1, since 00:28:33, 1920 weight, 1496 frag size
```

Frame Relay

Recommendation: For the latest feature combinations and management options, use class-based Frame Relay traffic shaping whenever possible.

Frame Relay networks are the most popular WANs in use today because of the low costs associated with them. Frame Relay is a nonbroadcast multiaccess (NBMA) technology that frequently utilizes oversubscription to achieve cost savings (similar to airlines overselling seats on flights to achieve maximum capacity and profitability).

To manage oversubscription and potential speed mismatches between senders and receivers, a traffic-shaping mechanism must be used with Frame Relay. Either Frame Relay traffic shaping (FRTS) or class-based FRTS can be used. The primary advantage of using class-based FRTS is management because shaping statistics and queuing statistics are displayed jointly with the **show policy interface** verification command and are included in the SNMPv2 Cisco class-based QoS Management Information Base (MIB).

FRTS and class-based FRTS require the following parameters to be defined:

- Committed information rate (CIR)
- Committed burst rate (Bc)
- Excess burst rate (Be)
- Minimum CIR
- Fragment size (required only on slow-speed links)

Committed Information Rate

Recommendation: Set the CIR to 95 percent of the PVC contracted speed.

In most Frame Relay networks, a central site's high-speed links connect to lower-speed links to/from many remote offices. For example, consider a central site that sends out data at 1.536 Mbps, while a remote branch might have only a 56-kbps circuit into it. This speed mismatch can cause congestion delays and drops. In addition, there is typically a many-to-one ratio of remote branches to central hubs, making it possible for many remote sites to send traffic at a rate that can overwhelm the T1 at the hub. Both scenarios can cause frame buffering in the provider network, which introduces jitter, delay, and loss.

The only solution to guarantee service-level quality is to use traffic shaping at both the central and remote routers and to define a consistent CIR at both ends of the Frame Relay PVC. Because the FRTS mechanism does not take Frame Relay overhead (headers and cyclic redundancy checks [CRCs]) into account in its calculations, it is recommended that the CIR be set slightly below the contracted speed of

the PVC. Cisco Technical Marketing testing has shown that setting the CIR to 95 percent of the contracted speed of the PVC engages the queuing mechanism (LLQ/CBWFQ) slightly early and improves service levels for Real-Time applications, like Voice.

Committed Burst Rate

Recommendation: Set the Bc to CIR/100 on nondistributed platforms and to CIR/125 on distributed platforms.

With Frame Relay networks, you also need to consider the amount of data that a node can transmit at any given time. A 56-kbps PVC can transmit a maximum of 56 kbps of traffic in 1 second. Traffic is not sent during the entire second, however, but only during a defined window called the interval (Tc). The amount of traffic that a node can transmit during this interval is called the committed burst (Bc) rate. By default, Cisco IOS Software sets the Bc to CIR/8. This formula is used for calculating the Tc follows:

Tc = Bc / CIR

For example, a CIR of 56 kbps is given a default Tc of 125 ms (7000 / 56,000). If the 56-kbps CIR is provisioned on a WAN aggregator that has a T1 line-rate clock speed, every time the router sends its allocated 7000 bits, it has to wait 120.5 ms before sending the next batch of traffic. Although this is a good default value for data, it is a bad choice for voice.

By setting the Bc value to a much lower number, you can force the router to send less traffic per interval, but over more frequent intervals per second. This results in significant reduction in shaping delays.

The optimal configured value for Bc is CIR/100, which results in a 10-ms interval (Tc = B / CIR).

On distributed platforms, the Tc must be defined in 4-ms increments. The nearest multiple of 4 ms within the 10-ms target is 8 ms. This interval can be achieved by configuring the Bc to equal CIR/125.

Excess Burst Rate

Recommendation: Set the Be to 0.

If the router does not have enough traffic to send all of its Bc (1000 bits, for example), it can "credit" its account and send more traffic during a later interval. The maximum amount that can be credited to the router's traffic account is called the excess burst (Be) rate. The problem with Be in converged networks is that this can create a potential for buffering delays within a Frame Relay network (because the receiving side can "pull" the traffic from a circuit only at the rate of Bc, not Bc + Be). To remove this potential for buffering delays, it is recommended to set the Be to 0.

Minimum Committed Information Rate

Recommendation: Set the minCIR to CIR.

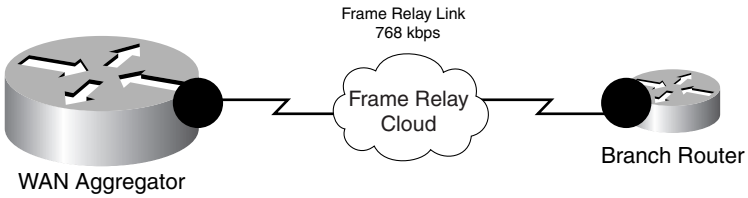
The minimum CIR is the transmit value that a Frame Relay router will "rate down" to when backward-explicit congestion notifications (BECNs) are received. By default, Cisco IOS Software sets the minimum CIR to CIR/2. However, to maintain consistent service levels, it is recommended that adaptive shaping be disabled and that the minimum CIR be set equal to the CIR (which means there is no "rating down"). An exception to this rule would occur if a tool such as Frame Relay voice-adaptive traffic shaping was deployed.

Slow-Speed (£ 768 kbps) Frame Relay Links

Recommendation: Enable FRF.12 and set the fragment size for 10 ms maximum serialization delay. Enable cRTP.

As with all slow-speed links, slow Frame Relay links (as illustrated in Figure 3-10) require a mechanism for fragmentation and interleaving. In the Frame Relay environment, the tool for accomplishing this is FRF.12.

Figure 3-10 Slow-Speed Frame Relay Links



Unlike MLP LFI, which takes the maximum serialization delay as a parameter, FRF.12 requires the actual fragment sizes to be defined manually. This requires some additional calculations because the maximum fragment sizes vary by PVC speed. These fragment sizes can be calculated by multiplying the provisioned PVC speed by the recommended maximum serialization delay target (10 ms), and converting the result from bits to bytes (which is done by dividing the result by 8):

Fragment Size in Bytes = (PVC Speed in kbps * Maximum Allowed Jitter in ms) / 8

For example, the calculation for the maximum fragment size for a 56-kbps circuit is as follows:

Fragment Size = (56 kbps * 10 ms) / 8 = 70 Bytes

Table 3-1 shows the recommended values for FRF.12 fragment sizes, CIR, and Bc for slow-speed Frame Relay links.

Table 3-1 Recommended Fragment Sizes, CIR, and Bc Values for Slow-Speed Frame Relay Links

PVC Speed	Maximum Fragment Size (for 10-ms Delay)	Recommended CIR Values	Recommended Bc Values
56 kbps	70 bytes	53,200 bps	532 bits per Tc
64 kbps	80 bytes	60,800 bps	608 bits per Tc
128 kbps	160 bytes	121,600 bps	1216 bits per Tc
256 kbps	320 bytes	243,200 bps	2432 bits per Tc
512 kbps	640 bytes	486,400 bps	4864 bits per Tc
768 kbps	960 bytes	729,600 bps	7296 bits per Tc

Both FRTS and class-based FRTS require a Frame Relay map class to be applied to the DLCI. Also in both cases, the **frame-relay fragment** command is applied to the map class. However, unlike FRTS, class-based FRTS does not require **frame-relay traffic-shaping** to be enabled on the main interface. This is because MQC-based/class-based FRTS requires a hierarchal (or nested) QoS policy to accomplish both shaping and queuing. This hierarchical policy is attached to the Frame Relay map class, which is bound to the DLCI.

As with slow-speed leased-line policies, cRTP can be enabled within the MQC queuing policy under the Voice class. Example 3-17 shows an example of slow-speed Frame Relay link-specific configuration.

Example 3-17 Slow-Speed (£ 768 kbps) Frame Relay QoS Design Example

```
!
policy-map MQC-FRTS-768
  class class-default
   shape average 729600 7296 0      ! Enables MQC-Based FRTS
   service-policy WAN-EDGE          ! Queues packets headed to the shaper

!
…
!
interface Serial2/0
 no ip address
 encapsulation frame-relay
!
interface Serial2/0.12 point-to-point
 ip address 10.1.121.1 255.255.255.252
 description 768kbps FR Circuit to RBR-3745-Left
 frame-relay interface-dlci 102
  class FR-MAP-CLASS-768              ! Binds the map-class to the FR DLCI
!
…
!
map-class frame-relay FR-MAP-CLASS-768
 service-policy output MQC-FRTS-768   ! Attaches nested MQC policies to map-class
 frame-relay fragment 960             ! Enables FRF.12
!
```

Verification commands:

- **show policy map**
- **show policy-map interface**
- **show frame-relay fragment**

Verification Command: **show frame-relay fragment**

The **show frame-relay fragment** command, shown in Example 3-18, provides verification of the fragment size, regardless of whether regular FRF.12 fragmentation or Frame Relay voice-adaptive traffic shaping (and fragmentation) is configured for a DLCI. Additionally, dynamic counters monitor how many frames required fragmentation in either direction.

*Example 3-18 **show frame-relay fragment** Verification of a Slow-Speed Frame Relay Link*

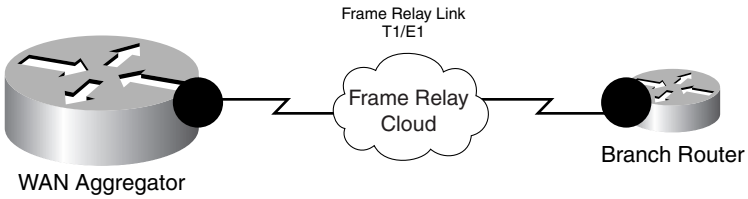
```
WAG-7206-Left#show frame-relay fragment 102
interface         dlci  frag-type    frag-size  in-frag    out-frag dropped-frag
Serial2/0.12      102   end-to-end   960        5476       2035       0
WAG-7206-Left#
```

Medium-Speed (£ T1/E1) Frame Relay Links

Recommendation: FRF.12 is not required. cRTP is optional.

The configuration for medium-speed Frame Relay links, illustrated in Figure 3-11 and detailed in Example 3-19, is identical to that for slow-speed Frame Relay links, with the exception that enabling FRF.12 no longer is required.

Figure 3-11 Medium-Speed Frame Relay Links



Note In some cases, however, administrators have chosen to enable FRF.12 on T1/E1 speed links, even though the fragment size for a 10-ms maximum serialization delay at such speeds is greater than the MTU of Ethernet (1500 bytes). The rationale behind doing so is to retain the Frame Relay dual-FIFO queuing mechanism at Layer 2, which can provide slightly superior service levels under certain conditions. Generally, this is not required however.

Example 3-19 Medium-Speed (T1/E1) Frame Relay QoS Design Example

```
!
policy-map MQC-FRTS-1536
  class class-default
   shape average 1460000 14600 0      ! Enables MQC-Based FRTS
   service-policy WAN-EDGE            ! Queues packets headed to the shaper
!
…
!
interface Serial2/0
 no ip address
 encapsulation frame-relay
!
interface Serial2/0.12 point-to-point
 ip address 10.1.121.1 255.255.255.252
 description 1536kbps FR Circuit to RBR-3745-Left
 frame-relay interface-dlci 102
  class FR-MAP-CLASS-1536           ! Binds the map-class to the FR DLCI
!
…
!
map-class frame-relay FR-MAP-CLASS-1536
 service-policy output MQC-FRTS-1536  ! Attaches nested MQC policies to map-class
!
```

Verification commands:

show policy map

show policy-map interface

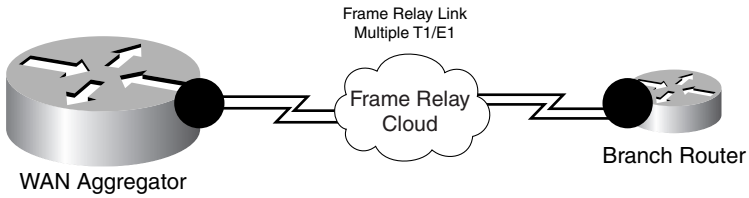
High-Speed (Multiple T1/E1 and Greater) Frame Relay Links

Recommendation: Use IP CEF per-packet load balancing for load sharing across multiple physical Frame Relay links.

When multiple Frame Relay circuits exist between a central WAN aggregation router and a remote branch router, as illustrated in Figure 3-12, it is recommended that IP CEF per-packet load balancing be used to load-share between the links. Multilink PPP over Frame Relay (MLPoFR) bundles are complex

to configure and difficult to manage, whereas IP CEF per-packet load balancing is not and has the lowest CPU impact of the load-sharing mechanisms. Therefore, IP CEF per-packet load balancing is recommended across multiple Frame Relay links to the same branch.

Figure 3-12 High-Speed Frame Relay Links



Note It is important to keep in mind that providers might have geographically dispersed paths to the same sites; therefore, the delay on one T1 FR link might be slightly higher or lower than the delay on another. This could cause TCP sequencing issues and slightly reduce effective data application throughput. Network administrators should keep these factors in mind when planning their WAN topologies.

The **max-reserved-bandwidth 100** command is not required on the interfaces because the queuing policy is not applied directly to the interface; instead, it is applied to another policy (the MQC-based Frame Relay traffic-shaping policy). Example 3-20 shows the configuration for a high-speed Frame Relay link.

Example 3-20 High-Speed (Š Multiple T1/E1) Frame Relay QoS Design Example

```
!
policy-map MQC-FRTS-1536
  class class-default
   shape average 1460000 14600 0    ! Enables MQC-Based FRTS
   service-policy WAN-EDGE          ! Queues packets headed to the shaper
!
…
!
interface Serial2/0
 no ip address
 encapsulation frame-relay
 no fair-queue
 frame-relay traffic-shaping
!
interface Serial2/0.12 point-to-point
 description 1536kbps FR Circuit to RBR-3745-Left
 ip address 10.1.121.1 255.255.255.252
 ip load-sharing per-packet        ! Enables IP CEF Per-Packet Load-Sharing
 frame-relay interface-dlci 102
  class FR-MAP-CLASS-1536          ! Binds the map-class to FR DLCI 102
!
interface Serial2/1
 no ip address
 encapsulation frame-relay
 serial restart_delay 0
!
interface Serial2/1.12 point-to-point
 description 1536kbps FR Circuit to RBR-3745-Left
 ip address 10.1.121.5 255.255.255.252
 ip load-sharing per-packet         ! Enables IP CEF Per-Packet Load-Sharing
 frame-relay interface-dlci 112
  class FR-MAP-CLASS-1536           ! Binds the map-class to FR DLCI 112
```

```
!
…
!
map-class frame-relay FR-MAP-CLASS-1536
 service-policy output MQC-FRTS-1536 ! Attaches nested MQC policies to map-class
!
```

Verification commands:

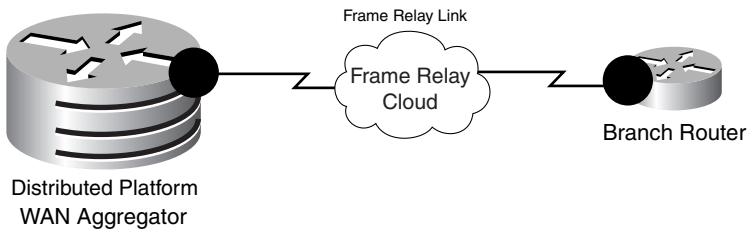
- **show policy map**
- **show policy-map interface**

Distributed Platform Frame Relay Links

Recommendation: Set CIR values to multiples of 8000. Set the Bc to CIR/125.

When ported to distributed-platform WAN aggregators (such as the Cisco 7500 VIP), most policies require little more than ensuring that IP CEF is running in distributed mode. However, FRTS is not supported in a distributed environment, so another shaping tool is required. Distributed traffic shaping (dTS) can be used in conjunction with hierarchical MQC policies to achieve a similar effect on traffic flows over distributed Frame Relay WAN links. Figure 3-13 shows a Frame Relay link homed from a distributed-platform WAN aggregator.

Figure 3-13 Distributed-Platform Frame Relay Links



Although dTS on the Cisco 7500 is very similar to MQC-based FRTS on nondistributed platforms, there are two main caveats to keep in mind. The first is that the CIR must be defined in multiples of 8000. Therefore, it is recommended that the CIR be defined as 95 percent of the PVC speed, rounded down to the nearest multiple of 8000. The second caveat is that the Cisco 7500 VIP requires the Tc to be defined in an increment of 4 ms. Because the target interval for all platforms is 10 ms, which is not evenly divisible by 4 ms, the recommendation for the Cisco 7500 VIP is to use an interval of 8 ms. The interval can be set to 8 ms by defining the burst using the following formula:

Bc = CIR/125

Table 3-2 gives recommended values for fragment sizes, CIR, and Bc for distributed platforms. (Some values have been slightly rounded for configuration and management simplicity.)

Table 3-2 Recommended Fragment Sizes, CIR, and Bc Values for Distributed Platform Frame Relay Links

PVC Speed	Maximum Fragment Size (for 10-ms Delay)	Recommended CIR Values	Recommended Bc Values
56 kbps	70 bytes	48,000 bps	384 bits per Tc
64 kbps	80 bytes	56,000 bps	448 bits per Tc
128 kbps	160 bytes	120,000 bps	960 bits per Tc

Table 3-2 Recommended Fragment Sizes, CIR, and Bc Values for Distributed Platform Frame Relay Links (continued)

PVC Speed	Maximum Fragment Size (for 10-ms Delay)	Recommended CIR Values	Recommended Bc Values
256 kbps	320 bytes	240,000 bps	1920 bits per Tc
512 kbps	640 bytes	480,000 bps	3840 bits per Tc
768 kbps	960 bytes	720,000 bps	5760 bits per Tc
1536 kbps	—	1,440,000 bps	11520 bits per Tc
2048 kbps	—	1,920,000 bps	15360 bits per Tc

Example 3-21 Distributed Platform Frame Relay QoS Design (Slow-Speed Link) Example

```
!
!
ip cef distributed       ! 'distributed' keyword required on 7500 for ip cef
!
…
!
policy-map MQC-DTS-768
  class class-default
    shape average 720000 5760 0      ! Enables Distributed Traffic Shaping
    service-policy WAN-EDGE          ! Queues packets headed to the shaper
!
…
!
interface Serial1/1/0
 no ip address
 encapsulation frame-relay
 no fair-queue
!
interface Serial1/1/0.12 point-to-point
 description 768kbps FR DLCI to RBR-3745-Left
 ip address 10.1.121.1 255.255.255.252
 frame-relay interface-dlci 102
  class FR-MAP-CLASS-768                 ! Binds the map-class to the FR-DLCI
!
…
!
map-class frame-relay FR-MAP-CLASS-768
 service-policy output MQC-DTS-768     ! Attaches nested MQC policies to map-class
 frame-relay fragment 960              ! Enables FRF.12
!
```

Verification commands:

- **show policy map**
- **show policy-map interface**
- **show frame-relay fragment** (on slow-speed links only)

ATM

As with Frame Relay, ATM is an NBMA medium that permits oversubscription and speed mismatches, and thus requires shaping to guarantee service levels. In ATM, however, shaping is included as part of the PVC definition.

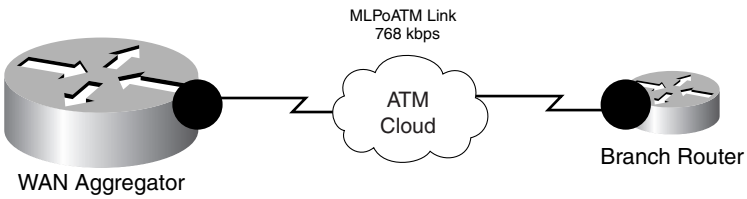
Two options exist for carrying voice traffic over slow-speed ATM PVCs: either Multilink PPP over ATM (MLPoATM), in conjunction with MLP LFI, or ATM PVC bundling. ATM PVC bundling is a legacy technique that has drawbacks such as inefficient bandwidth utilization and classification limitations (IP precedence versus DSCP). But sometimes service providers make ATM PVC bundles economically attractive to enterprise customers, so both approaches are discussed.

Slow-Speed (£ 768 kbps) ATM Links: MLPoATM

Recommendation: Use MLP LFI. Tune the ATM PVC Tx-ring to 3. cRTP can be used only in Cisco IOS Release 12.2(2)T or later.

Serialization delays on slow-speed ATM links, as shown in Figure 3-14, necessitate a fragmentation and interleaving mechanism. The most common ATM adaptation layers (such as AAL5) do not have sequence numbers in the cell headers and, thus, require cells to arrive in the correct order. This requirement makes interleaving a problem that cannot be solved at these ATM adaptation layers and thus must be solved at a higher layer.

Figure 3-14 Slow-Speed MLPoATM Links



A solution to this problem is to run MLPoATM and let MLP LFI handle any necessary fragmentation and interleaving so that such operations are completely transparent to the lower ATM layer. As far as the ATM layer is concerned, all cells arrive in the same order they were sent.

MLPoATM functionality is enabled through the use of virtual-access interfaces. Virtual-access interfaces are built on demand from virtual-template interfaces and inherit their configuration properties from the virtual templates they are built from. Thus, the IP address, **service-policy** statement, and LFI parameters all are configured on the virtual template, as shown in Example 3-22.

cRTP is supported only on ATM PVCs (through MLPoATM), as of Cisco IOS Release 12.2(2)T.

Additionally, as discussed previously in this chapter, it is recommended that the value of the final output buffer, the Tx-ring, be tuned on slow-speed ATM PVCs to a value of three particles to minimize serialization delay.

Example 3-22 Slow-Speed (£ 768 kbps) MLPoATM QoS Design Example

```
!
interface ATM4/0
 bandwidth 768
 no ip address
 no atm ilmi-keepalive
!
interface ATM4/0.60 point-to-point
 pvc BRANCH#60 0/60
  vbr-nrt 768 768                    ! ATM PVC definition
  tx-ring-limit 3                    ! Per-PVC Tx-ring is tuned to 3 particles
  protocol ppp Virtual-Template60    ! PVC is bound to the Virtual-Template
!
interface Virtual-Template60
 bandwidth 768
```

```
 ip address 10.200.60.1 255.255.255.252
 service-policy output WAN-EDGE       ! Attaches MQC policy to Virtual-Template
 ppp multilink
 ppp multilink fragment-delay 10      ! Enables MLP Fragmentation
 ppp multilink interleave             ! Enables MLP Interleaving
!
```

Note When using virtual templates for low-speed ATM links, keep the following in mind:

-- The dynamic nature of virtual-template interfaces might make network management unwieldy.

-- MLPoATM can be supported only on hardware that supports per-VC traffic shaping.

Verification commands:

- show policy map
- show policy-map interface
- show atm pvc

Verification Command: show atm pvc

In ATM, the length of the Tx-ring is defined in particles, not packets. The size of a particle varies according to hardware. For example, on a Cisco 7200 PA-A3, particles are 580 bytes (including a 4-byte ATM core header). This means that a 1500-byte packet would require three particles of buffering. Furthermore, ATM defines Tx-rings on a per-PVC basis, as shown in Example 3-23 and Example 3-24.

Example 3-23 Basic ATM PVC Configuration Example

```
!
interface ATM3/0.1 point-to-point
 ip address 10.2.12.1 255.255.255.252
 pvc 0/12
  vbr-nrt 768 768     ! ATM PVC definition
 !
!
```

The size of a default Tx-ring can be ascertained using the **show atm pvc** command (an output modifier is used to focus on the relevant portion of the output), as shown in Example 3-24.

Example 3-24 show atm pvc Verification of Tx-ring Setting

```
WAG-7206-Left#show atm pvc 0/12 | include TxRingLimit
VC TxRingLimit: 40 particles
```

The output shows that the Tx-ring is set, in this instance, to a default value of 40 particles. The Tx-ring for the PVC can be tuned to the recommended setting of 3 using the **tx-ring-limit** command under the PVC's definition, as shown in Example 3-25.

Example 3-25 Tuning an ATM PVC Tx-ring

```
WAG-7206-Left(config)#interface atm 3/0.1
WAG-7206-Left(config-subif)#pvc 0/12
WAG-7206-Le(config-if-atm-vc)#tx-ring-limit 3
```

The new setting can be verified quickly with the same **show atm pvc** command variation, as shown in Example 3-25 (see Example 3-26).

Example 3-26 show atm pvc Verification of Tx-ring Setting After Tuning

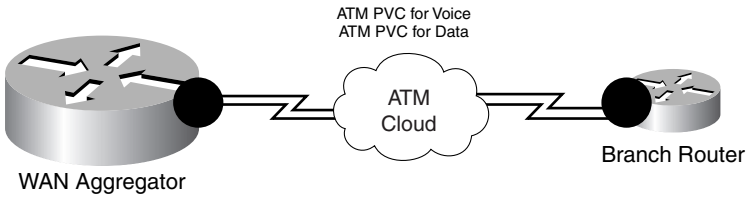
```
WAG-7206-Left#show atm pvc 0/12 | include TxRingLimit
VC TxRingLimit: 3 particles
```

Slow-Speed (£ 768 kbps) ATM Links: ATM PVC Bundles

Recommendation: Queuing policies for voice are not required (because voice uses a dedicated ATM PVC). Tune the ATM PVC Tx-ring to 3.

An alternative option to provisioning QoS on slow-speed ATM PVCs is to use PVC bundles, as illustrated in Figure 3-15. PVC bundles consist of two (or more) PVCs with different ATM traffic contracts, grouped together in a logical association in which IPP levels determine the PVC to which the packet will be directed. The decision to use PVC bundles instead of MLPoATM for slow-speed ATM links is usually a matter of economics (because service providers often offer attractive pricing for PVC bundles) and configuration/management complexity comfort levels.

Figure 3-15 Slow-Speed ATM PVC Bundles



In Example 3-27, one PVC (for voice) has a variable bit rate, non-real-time (VBR-nrt) ATM traffic contract and an admission criterion of IPP 5, while another PVC (for data) has an unspecified bit rate (UBR) ATM traffic contract and accepts all other precedence levels.

Again, it is also recommended that the TX-ring be tuned to 3 on such slow-speed ATM PVCs.

Example 3-27 Slow-Speed (ð 768 kbps) ATM PVC Bundles QoS Design Example

```
!
class-map match-any Call Signaling
  match ip dscp cs3
match ip dscp af31
class-map match-any Critical Data
  match ip dscp cs6
  match ip dscp af21
  match ip dscp cs2
!
!
 policy-map WAN-EDGE-DATA-PVC     ! Only data queuing is required (no voice)
  class Call Signaling
   bandwidth percent 5
  class Critical Data
   bandwidth percent 40
  class class-default
   fair-queue
!
vc-class atm VOICE-PVC-256     ! Voice PVC-class definition
```

```
  vbr-nrt 256 256              ! Voice ATM PVC definition
  tx-ring-limit 3              ! Per-PVC Tx-ring is tuned to 3 particles
  precedence 5                 ! Only IPP5 traffic (voice) can use this PVC
  no bump traffic              ! Traffic will not be accepted from other PVCs
  protect vc                   ! Optional: Protects VC status of Voice PVC

!
vc-class atm DATA-PVC-512     ! Data PVC-class definition
  ubr 512                     ! Data ATM PVC definition
  tx-ring-limit 3             ! Per-PVC Tx-ring is tuned to 3 particles
  precedence other            ! All other IPP values (data) use this PVC
!
…
!
interface ATM3/0
 no ip address
 no atm ilmi-keepalive
!
interface ATM3/0.60 point-to-point
 ip address 10.200.60.1 255.255.255.252
 bundle BRANCH#60
  pvc-bundle BRANCH60-DATA 0/60
   class-vc DATA-PVC-512                      ! Assigns PVC to data-class
   service-policy output WAN-EDGE-DATA-PVC    ! Attaches (data) MQC policy to PVC
  pvc-bundle BRANCH60-VOICE 0/600
   class-vc VOICE-PVC-256                     ! Assigns PVC to voice-class
!
```

A major drawback to PVC bundling is that data never can get access to the voice PVC, even if there is available bandwidth in it. This forces suboptimal consumption of WAN bandwidth.

Verification commands:

- **show policy map**
- **show policy-map interface**
- **show atm pvc**
- **show atm vc**
- **show atm bundle**

Verification Command: **show atm vc**

The **show atm vc** command details the configured ATM PVCs and highlights their encapsulation, ATM traffic contracts (or service contracts), status, and activity, as shown in Example 3-28.

Example 3-28 **show atm vc** ***Verification of ATM PVC Definitions and Activity***

```
WAN-AGG-7200#show atm vc
           VCD /                                        Peak  Avg/Min Burst
Interface  Name          VPI VCI Type   Encaps   SC    Kbps   Kbps   Cells  Sts
3/0.60     BRANCH60-DATA  0  60 PVC    SNAP     UBR    512    512    1145   UP
3/0.60     BRANCH60-VOICE 0 600 PVC    SNAP     VBR    256    256     94    UP
WAN-AGG-7200#
```

Verification Command: show atm bundle

The **show atm bundle** command provides details on the configured and current admission criteria for individual ATM PVCs. In Example 3-29, PVC 0/600 (the voice PVC) accepts only traffic that has been marked to IPP 5 (voice). All other IPP values (0 to 4 and 6 to 7) are assigned to PVC 0/60 (the data PVC). This command also shows the activity for each PVC.

Example 3-29 show atm bundle Verification of ATM PVC Bundle Definitions and Activity

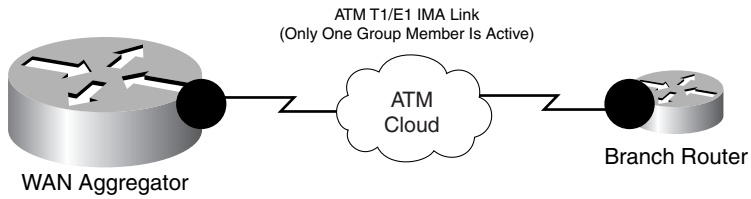
```
WAN-AGG-7200#show atm bundle
BRANCH#60 on ATM3/0.60: UP
                      Config    Current   Bumping  PG/ Peak Avg/Min Burst
VC Name       VPI/ VCI Prec/Exp  Prec/Exp  PrecExp/ PV  Kbps   kbps  Cells Sts
                                           Accept
BRANCH60-DATA  0/60    7-6, 4-0  7-6, 4-0  - / Yes  -     512  512   1145   UP
BRANCH60-VOICE 0/600   5         5         - / No   PV    256  256    94    UP
WAN-AGG-7200#
```

Medium-Speed (£ T1/E1) ATM Links

Recommendation: Use ATM inverse multiplexing over ATM (IMA) to keep future expansion easy to manage. No LFI is required. cRTP is optional.

ATM IMA is a natural choice for medium-speed ATM links, as shown in Figure 3-16. Although the inverse-multiplexing capabilities are not used at these speeds, IMA interfaces make future expansion to high-speed links easy to manage (as will be demonstrated between Example 3-30 and the high-speed ATM link in Example 3-35).

Figure 3-16 Medium-Speed ATM Links



Example 3-30 Medium-Speed (T1/E1) ATM IMA QoS Design Example

```
!
interface ATM3/0
 no ip address
 no atm ilmi-keepalive
 ima-group 0            ! ATM3/0 added to ATM IMA group 0
 no scrambling-payload

!
…
!
interface ATM3/IMA0
 no ip address
 no atm ilmi-keepalive
!
interface ATM3/IMA0.12 point-to-point
 ip address 10.200.60.1 255.255.255.252
```

```
 description T1 ATM-IMA to Branch#60
 pvc 0/100
  vbr-nrt 1536 1536               ! ATM PVC defined under ATM IMA sub-int
  max-reserved-bandwidth 100      ! Overrides the default 75% BW limit
  service-policy output WAN-EDGE  ! Attaches MQC policy to PVC
 !
!
```

Verification commands:

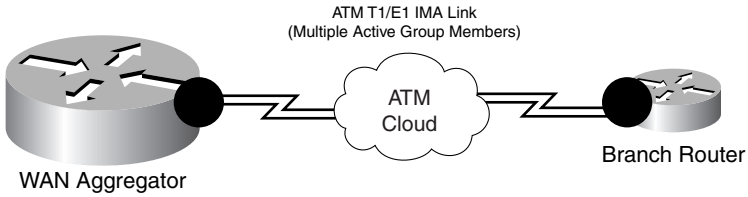
- **show policy map**
- **show policy-map interface**
- **show atm pvc**
- **show ima interface atm**

High-Speed (Multiple T1/E1) ATM Links

Recommendation: Use ATM IMA and add members to the IMA group, as needed.

Previous options for accommodating multiple T1/E1 links were software-based load-sharing solutions (MLP bundling and IP CEF per-packet load sharing). As such, these methods require additional CPU cycles to accommodate load-sharing multiple physical links. However, with ATM IMA, inverse multiplexing over multiple T1/E1 links, illustrated in Figure 3-17, is done in hardware on the port adaptor/network module. Therefore, ATM IMA scales much more efficiently.

Figure 3-17 High-Speed ATM Links



As mentioned, ATM IMA makes bandwidth expansion easy to manage. For example, all that is required to add another T1 line to the previous example is to add an **ima-group** statement to the next ATM interface and increase the PVC speed, as shown in Example 3-31.

Example 3-31 High-Speed (Multiple T1/E1 and Greater) ATM IMA QoS Design Example

```
!
interface ATM3/0
 no ip address
 no atm ilmi-keepalive
 ima-group 0        ! ATM3/0 added to ATM IMA group 0
 no scrambling-payload
!
interface ATM3/1
 no ip address
 no atm ilmi-keepalive
 ima-group 0        ! ATM3/1 added to ATM IMA group 0
 no scrambling-payload
!
…
!
interface ATM3/IMA0
 no ip address
```

```
 no atm ilmi-keepalive
!
interface ATM3/IMA0.12 point-to-point
 ip address 10.6.12.1 255.255.255.252
 pvc 0/100
  vbr-nrt 3072 3072                  ! ATM PVC speed expanded
  max-reserved-bandwidth 100         ! Overrides the default 75% BW limit
  service-policy output WAN-EDGE     ! Attaches MQC policy to PVC
 !
!
```

Verification commands:

- **show policy map**
- **show policy-map interface**
- **show atm pvc**
- **show ima interface atm**

Verification Command: show ima interface atm

The **show ima interface atm** command is useful for verifying that all members of an ATM IMA group are active. See Example 3-32.

Example 3-32 **show ima interface atm** ***Verification of ATM IMA Group***

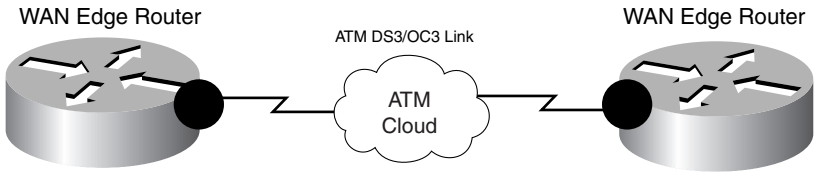
```
WAG-7206-Left#show ima interface atm 3/ima0
Interface ATM3/IMA0 is up
        Group index is 1
        Ne state is operational, failure status is noFailure
        Active links bitmap 0x3
    IMA Group Current Configuration:
        Tx/Rx configured links bitmap 0x3/0x3
        Tx/Rx minimum required links 1/1
        Maximum allowed diff delay is 25ms, Tx frame length 128
        Ne Tx clock mode CTC, configured timing reference link ATM3/0
        Test pattern procedure is disabled
    IMA Group Current Counters (time elapsed 257 seconds):
        0 Ne Failures, 0 Fe Failures, 0 Unavail Secs
    IMA Group Total Counters (last 5 15 minute intervals):
        0 Ne Failures, 0 Fe Failures, 0 Unavail Secs
    IMA link Information:
        Link     Physical Status        NearEnd Rx Status      Test Status
        ----     ---------------        -----------------      -----------
        ATM3/0   up                     active                 disabled
        ATM3/1   up                     active                 disabled
        ATM3/2   administratively down  unusableInhibited      disabled
        ATM3/3   administratively down  unusableInhibited      disabled
```

Very-High-Speed (DS3-OC3+) ATM Links

Recommendation: Use newer hardware platforms and keep an eye on CPU levels.

Major site-to-site interconnections drift slightly away from the traditional WAN aggregator/remote branch router models. In site-to-site scenarios, as illustrated in Figure 3-18, the WAN edge routers usually support only one or two links, as opposed to dozens or hundreds of links that typical WAN aggregators support. However, in a site-to-site scenario, the interconnecting links are running at far higher speeds than most remote branch links.

Figure 3-18 Very High-Speed (DS3-OC3+) ATM Links



The policies and design principles do not change for site-to-site scenarios. The main consideration is the performance of the WAN edge router. Although newer platforms handle complex policies more efficiently, it is still highly recommended that proof-of-concept testing of the platforms involved be performed before implementing policies at such critical junctions in the network. Example 3-33 illustrates a site-to-site QoS policy applied to a very-high-speed ATM (OC3) link.

Example 3-33 Very High-Speed (DS3-OC3+) ATM Link QoS Design Example

```
!
interface ATM3/0
 no ip address
 load-interval 30
 no atm ilmi-keepalive
!
interface ATM3/0.1 point-to-point
 ip address 10.2.12.1 255.255.255.252
 pvc 0/12
  vbr-nrt 149760 149760              ! ATM OC3 PVC definition
  max-reserved-bandwidth 100         ! Overrides the default 75% BW limit
  service-policy output WAN-EDGE     ! Attaches MQC policy to PVC
 !
!
```

Verification commands:

- show policy map
- show policy-map interface
- show atm pvc

ATM-to-Frame Relay Service Interworking

Many enterprises are deploying converged networks that use ATM at the central site and Frame Relay at the remote branches. The media conversion is accomplished through ATM-to-Frame Relay Service Interworking (SIW or FRF.8) in the carrier network.

FRF.12 cannot be used because, currently, no service provider supports FRF.12 termination in the Frame Relay cloud. In fact, no Cisco WAN switching devices support FRF.12. Tunneling FRF.12 through the service provider's network does no good because there is no FRF.12 standard on the ATM side. This is

a problem because fragmentation is a requirement if any of the remote Frame Relay sites uses a circuit speed of 768 kbps or below. However, MLPoATM and MLPoFR provide an end-to-end, Layer 2 fragmentation and interleaving method for low-speed ATM to Frame Relay FRF.8 SIW links.

FRF.8 SIW is a Frame Relay Forum standard for connecting Frame Relay networks with ATM networks. SIW provides a standards-based solution for service providers, enterprises, and end users. In service interworking translation mode, Frame Relay PVCs are mapped to ATM PVCs without the need for symmetric topologies. FRF.8 supports two modes of operation of the interworking function (IWF) for upper-layer user protocol encapsulation:

- **Translation mode**—Maps between ATM (AAL) and Frame Relay (IETF) encapsulation. It also supports interworking of routed or bridged protocols.
- **Transparent mode**—Does not map encapsulations, but sends them unaltered. This mode is used when translation is impractical because encapsulation methods do not conform to the supported standards for service interworking.

MLP for LFI on ATM and Frame Relay SIW networks is supported for transparent-mode VCs and translational-mode VCs that support PPP translation (FRF 8.1).

To make MLPoATM and MLPoFR SIW possible, the service provider's interworking switch must be configured in transparent mode, and the end routers must be capable of recognizing both MLPoATM and MLPoFR headers. This is accomplished with the **protocol ppp** command for ATM and the **frame-relay interface-dlci** *dlci* **ppp** command for Frame Relay.

When an ATM cell is sent from the ATM side of an ATM-to-Frame Relay SIW connection, the following must happen for interworking to be possible:

1. The sending router encapsulates a packet in the MLPoATM header by the sending router.
2. In transparent mode, the carrier switch prepends a 2-byte Frame Relay DLCI field to the received packet and sends the packet to its Frame Relay interface.
3. The receiving router examines the header of the received packet. If the first 4 bytes after the 2-byte DLCI field of the received packet are 0xfefe03cf, it treats it as a legal MLPoFR packet and sends it to the MLP layer for further processing.

When a frame is sent from the Frame Relay side of an ATM-to-Frame Relay SIW connection, the following must happen for interworking to be possible:

1. The sending router encapsulates a packet in the MLPoFR header.
2. In transparent mode, the carrier switch strips off the 2-byte Frame Relay DLCI field and sends the rest of the packet to its ATM interface.
3. The receiving router examines the header of the received packet. If the first 2 bytes of the received packet are 0x03cf, it treats it as a legal MLPoATM packet and sends it to MLP layer for further processing.

A new ATM-to-Frame Relay SIW standard, FRF.8.1, supports MLPoATM and Frame Relay SIW, but it could be years before all switches are updated to this new standard.

When using MLPoATM and MLPoFR, keep the following in mind:

- MLPoATM can be supported only on platforms that support per-VC traffic shaping.
- MLPoATM relies on per-VC queuing to control the flow of packets from the MLP bundle to the ATM PVC.
- MLPoATM requires the MLP bundle to classify the outgoing packets before they are sent to the ATM VC. It also requires the per-VC queuing strategy for the ATM VC to be FIFO because the MLP bundle handles queuing.
- MLPoFR relies on the FRTS engine to control the flow of packets from the MLP bundle to FR VC.

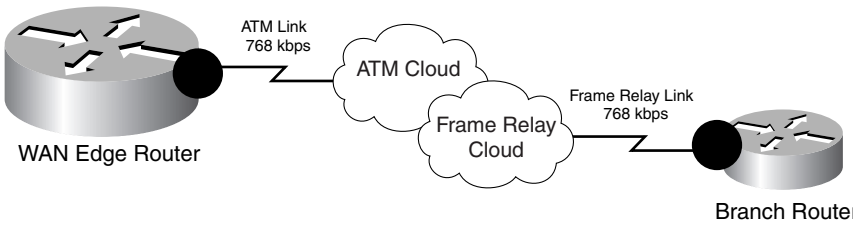
- cRTP is supported only over ATM links (through MLPoATM), as of Cisco IOS Release 12.2(2)T.

Slow-Speed (£ 768 kbps) ATM-FR SIW Links

Recommendation: Use MLPoATM and MLPoFR. Use MLP LFI and optimize fragment sizes to minimize cell padding. cRTP can be used only in Cisco IOS Release 12.2(2)T or later. Tune the ATM PVC Tx-ring to 3.

As with any slow-speed WAN media, serialization delay must be addressed with a fragmentation and interleaving mechanism. As previously mentioned, FRF.12 is not an option for SIW links. Therefore, MLP LFI must be used. Generally, MLP LFI requires no additional calculations to configure, but a special case exists when interworking ATM and FR (as illustrated in Figure 3-19) because of the nature of ATM's fixed cell lengths.

Figure 3-19 Slow-Speed ATM-FR SIW Links



When enabling MLPoATM, the fragment size should be optimized so that it fits into an integral number of cells. Otherwise, the bandwidth required could double because of cell padding. For example, if a fragment size of 49 bytes is configured, this fragment would require 2 cells to transmit (because ATM cells have 48-byte payloads). This would generate 57 bytes of overhead (2 cell headers plus 47 bytes of cell padding), which is more than double the fragment itself.

Table 3-3 provides a summary of the optimal fragment-delay parameters for MLPoATM.

Table 3-3 Optimal Fragment-Delay Values for MLP LFI for MLPoATM

PVC Speed	Optimal Fragment Size	ATM Cells (Rounded Up)	ppp multilink fragment-delay value
56 kbps	84 bytes	2	12 ms
64 kbps	80 bytes	2	10 ms
128 kbps	176 bytes	4	11 ms
256 kbps	320 bytes	7	10 ms
512 kbps	640 bytes	14	10 ms
768 kbps	960 bytes	21	10 ms

A slow-speed ATM-to-Frame Relay SIW configuration is shown next, in two parts:

- The central site WAN aggregator MLPoATM configuration (see Example 3-34)
- The remote branch router MLPoFR configuration (see Example 3-35)

Example 3-34 MLPoATM WAN Aggregator ATM-FR SIW QoS Design Example

```
!
interface ATM4/0
 no ip address
```

```
 no atm ilmi-keepalive
!
interface ATM4/0.60 point-to-point
 pvc BRANCH#60 0/60
  vbr-nrt 256 256                      ! ATM PVC definition
  tx-ring-limit 3                      ! Per-PVC
Tx-ring is tuned to 3 particles
  protocol ppp Virtual-      Template60     ! Enables MLPoATM
!
interface Virtual-Template60
 bandwidth 256
 ip address 10.200.60.1 255.255.255.252
 service-policy output WAN-EDGE       ! Attaches MQC policy to Virtual-Template
 ppp multilink
 ppp multilink fragment-delay 10      ! Enables MLP fragmentation
 ppp multilink interleave             ! Enables MLP interleaving
!
```

Verification commands:

- **show policy map**
- **show policy-map interface**
- **show atm pvc**
- **show ppp multilink**

Example 3-35 MLPoFR Remote-Branch Router ATM-FR SIW QoS Design Example

```
!
interface Serial6/0
 description Parent FR Link for BRANCH#60
 no ip address
 encapsulation frame-relay
 frame-relay traffic-shaping
!
interface Serial6/0.60 point-to-point
 description FR Sub-Interface for BRANCH#60
 bandwidth 256
 frame-relay interface-dlci 60 ppp Virtual-Template60      ! Enables MLPoFR
  class FRTS-256kbps                ! Binds the map-class to the FR DLCI
!
interface Virtual-Template60
 bandwidth 256
 ip address 10.200.60.2 255.255.255.252
 service-policy output WAN-EDGE      ! Attaches MQC policy to map-class
 ppp multilink
 ppp multilink fragment-delay 10     ! Enables MLP fragmentation
 ppp multilink interleave            ! Enables MLP interleaving
!
…
!
map-class frame-relay FRTS-256kbps
 frame-relay cir 243200              ! CIR is set to 95% of FR DLCI rate
 frame-relay bc 2432                 ! Bc is set to CIR/100
 frame-relay be 0                    ! Be is set to 0
 frame-relay mincir 243200           ! MinCIR is set to CIR
!
```

Verification commands:

- **show policy map**

- **show policy-map interface**
- **show ppp multilink**

ISDN

When designing VoIP over ISDN networks, special consideration needs to be given to the following issues:

- Link bandwidth varies as B channels are added or dropped.
- RTP packets might arrive out of order when transmitted across multiple B channels.
- CallManager has limitations with locations-based CAC.

Variable Bandwidth

ISDN allows B channels to be added or dropped in response to the demand for bandwidth. The fact that the bandwidth of a link varies over time presents a special challenge to the LLQ/CBWFQ mechanisms of Cisco IOS Software. Before Cisco IOS Release 12.2(2)T, a policy map implementing LLQ could be assigned only a fixed amount of bandwidth. On an ISDN interface, Cisco IOS Software assumes that only 64 kbps is available, even though the interface has the potential to provide 128 kbps, 1.544 Mbps, or 2.408 Mbps of bandwidth. By default, the maximum bandwidth assigned must be less than or equal to 75 percent of the available bandwidth. Hence, before Cisco IOS Release 12.2(2)T, only 75 percent of 64 kbps, or 48 kbps, could be allocated to an LLQ on any ISDN interface. If more was allocated, an error message was generated when the policy map was applied to the ISDN interface. This severely restricted the number of VoIP calls that could be carried.

The solution to this problem was introduced in Cisco IOS Release 12.2(2)T with the **priority percent** command. This command allows the reservation of a variable bandwidth percentage to be assigned to the LLQ.

MLP Packet Reordering Considerations

MLP LFI is used for fragmentation and interleaving voice and data over ISDN links. LFI segments large data packets into smaller fragments and transmits them in parallel across all the B channels in the bundle. At the same time, voice packets are interleaved between the fragments, thereby reducing their delay. The interleaved packets are not subject to MLP encapsulation; they are encapsulated as regular PPP packets. Hence, they have no MLP sequence numbers and cannot be reordered if they arrive out of sequence.

The packets probably will need to be reordered. The depth of the various link queues in the bundle might differ, causing RTP packets to overtake each other as a result of the difference in queuing delay. The various B channels also might take different paths through the ISDN network and might end up with different transmission delays.

This reordering of packets is not generally a problem for RTP packets. The buffers on the receiving VoIP devices reorder the packets based on the RTP sequence numbers. However, reordering becomes a problem if cRTP is used. The cRTP algorithm assumes that RTP packets are compressed and decompressed in the same order. If they get out of sequence, decompression does not occur correctly.

Multiclass Multilink PPP (MCMP) offers a solution to the reordering problem. With MCMP, the interleaved packets are given a small header with a sequence number, which allows them to be reordered by the far end of the bundle before cRTP decompression takes place. MCMP is supported as of Cisco IOS Release 12.2(13)T.

CallManager CAC Limitations

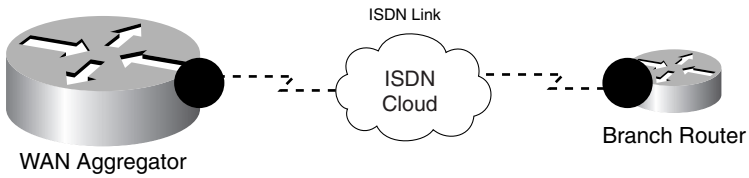
IP telephony in branch networks typically is based on the centralized call-processing model and uses locations-based CAC to limit the number of calls across the WAN. Locations-based CAC currently does not have any mechanism for tracking topology changes in the network. Therefore, if the primary link to a branch goes down and ISDN backup engages, the CallManager remains ignorant of the occurrence. For this reason, it is critical that the ISDN backup link be capable of handling the same number of VoIP calls as the main link. Otherwise, CAC ultimately could oversubscribe the backup link.

The actual bandwidth of the primary link and the backup link do not need to be identical. They just need to be capable of carrying the same number of VoIP calls. For example, the backup link might use cRTP while the primary link does not, in which case, less bandwidth is required on the backup link to carry the same number of calls as the primary link. Because of these limitations, it is recommended that the 33 percent LLQ recommendation be relaxed in this kind of dial-backup scenario. The LLQ could be provisioned as high as 70 percent (leaving 5 percent for Voice control traffic over the ISDN link and 25 percent for Best-Effort traffic).

Voice and Data on Multiple ISDN B Channels

The Voice and Data design model over ISDN, illustrated in Figure 3-20, allows a service policy to be applied to a bundle with multiple B channels. It takes advantage of the fact that LLQ bandwidth can be expressed as a percentage instead of an absolute number. If cRTP is enabled, MCMP is required on the ISDN links.

Figure 3-20 Voice and Data over ISDN



Cisco IOS provides two mechanisms for controlling how channels are added in response to demand.

The first mechanism commonly is referred to as dial-on-demand routing (DDR). With DDR, a load threshold must be specified (as a fraction of available bandwidth). When the traffic load exceeds this number, an additional channel is added to the bundle. The threshold is calculated as a running average. As a result, there is a certain delay in bringing up additional B channels when the load increases. This delay does not present a problem with data, but it is unacceptable with voice. This delay can be reduced to around 30 seconds by adding the **load-interval** command to the physical ISDN interface, but even 30 seconds is too long.

The second mechanism is a more robust solution, which is simply to bring up all B channels immediately and keep them up as long as the ISDN service is required. This is achieved by using the **ppp multilink links minimum** command.

With two B channels available, the service policy can reserve (approximately) 90 kbps (70 percent of 128 kbps) for voice traffic. The total number of calls that can be transmitted depends on the codec and sampling rates used.

Example 3-36 illustrates the configuration for enabling voice and data over multiple ISDN B channels.

Example 3-36 Voice and Data over Multiple ISDN B Channels QoS Design Example

```
!
class-map match-all Voice
  match ip dscp ef
!
class-map match-any Call Signaling
  match ip dscp cs3
  match ip dscp af31
!
…
!
policy-map WAN-EDGE-ISDN
  class Voice
    priority percent 70       ! LLQ 33% Rule is relaxed for DDR scenarios
    compress header ip rtp    ! Enables Class-Based cRTP
 class Call Signaling
   bandwidth percent 5        ! Bandwidth guarantee for Call-Signaling
  class class-default
   fair-queue
!
interface BRI0/0
 encapsulation ppp
 dialer pool-member 1
!
interface Dialer1
 encapsulation ppp
 dialer pool 1
 dialer remote-name routerB-dialer1
 dialer-group 1
 dialer string 12345678
 service-policy output WAN-EDGE-ISDN     ! Attaches MQC policy to Dialer interface
 ppp multilink
 ppp multilink fragment-delay 10         ! Enables MLP fragmentation
 ppp multilink interleave                ! Enables MLP interleaving
 ppp multilink links minimum 2           ! Activates both B Channels immediately
 ppp multilink multiclass                ! Enables MCMP
!
```

Verification commands:

- **show policy map**
- **show policy-map interface**
- **show ppp multilink**

Note For a case study example of WAN Aggregator QoS design, refer to Figure 13-21 and Example 13-37 of the Cisco Press book, *End-to-End QoS Network Design* by Tim Szigeti and Christina Hattingh.

Summary

This chapter discussed the QoS requirements of routers performing the role of a WAN aggregator. Specifically, it addressed the need for queuing policies on the WAN edges, combined with shaping policies when NBMA media (such as Frame Relay or ATM) are being used, and link-specific policies, such as LFI/FRF.12 and cRTP, for slow-speed (£ 768 kbps) links.

For the WAN edges, bandwidth-provisioning guidelines were considered, such as leaving 25 percent of the bandwidth for the Best-Effort class and limiting the sum of all LLQs to 33 percent.

Three categories of WAN link speeds and their design implications were presented:

- Slow-speed (£ 768 kbps) links, which can support only Three- to Five-Class QoS models and require LFI mechanisms and cRTP.
- Medium-speed (£ T1/E1) links, which, likewise, can support only Three- to Five-Class QoS Models but no longer require LFI mechanisms. cRTP becomes optional.
- High-speed (multiple T1/E1 or greater) links, which can support 5- to 11-Class QoS Models. No LFI is required on such links. cRTP likely would have a high CPU cost (compared to realized bandwidth savings) and, as such, generally is not recommended for such links. Additionally, some method of load sharing, bundling, or inverse multiplexing is required to distribute the traffic across multiple physical links.

These principles then were applied to certain WAN media designs—specifically, for leased lines, Frame Relay, ATM, and ATM-to-Frame Relay SIW. The corner case of ISDN as a backup WAN link also was considered.

References

Standards

- RFC 2474 “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers” http://www.ietf.org/rfc/rfc2474
- RFC 2597 “Assured Forwarding PHB Group” http://www.ietf.org/rfc/rfc2597
- RFC 2697 “A Single Rate Three Color Marker” http://www.ietf.org/rfc/rfc2697
- RFC 2698 “A Two Rate Three Color Marker” http://www.ietf.org/rfc/rfc2698
- RFC 3168 “The Addition of Explicit Congestion Notification (ECN) to IP” http://www.ietf.org/rfc/rfc3168
- RFC 3246 “An Expedited Forwarding PHB (Per-Hop Behavior)” http://www.ietf.org/rfc/rfc3246

Books

- Szigeti, Tim and Christina Hattingh. *End-to-End QoS Network Design: Quality of Service in LANs, WANs and VPNs*. Indianapolis: Cisco Press, 2004.

Cisco Documentation

Layer 3 queuing:

- Class-based weighted fair queuing (Cisco IOS Release 12.0.5T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t5/cbwfq.htm

- Low-latency queuing (Cisco IOS Release 12.0.7T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t7/pqcbwfq.ht
- Distributed low-latency queuing (Cisco IOS Release 12.1.5T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t5/dtllqvip.htm
- Low-latency queuing with priority percentage support (Cisco IOS Release 12.2.2T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t2/ftllqpct.htm

Congestion avoidance:

- MQC-based WRED (Cisco IOS Release 12.0.5T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t5/cbwfq.htm
- DiffServ-compliant weighted random early detection (Cisco IOS Release 12.1.5T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t5/dtdswred.htm
- Distributed class-based weighted fair queuing and distributed weighted random early detection (Cisco IOS Release 12.1.5T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t5/dtcbwred.htm

Frame Relay traffic shaping:

- Class-based shaping (Cisco IOS Release 12.1.2T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t2/clsbsshp.htm
- MQC-based Frame Relay traffic shaping (Cisco IOS Release 12.2.13T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/frqosmqc.htm
- Distributed traffic shaping (Cisco IOS Release 12.1.5T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t5/dtdts.htm

ATM PVC trafficparameters:

- Configuring ATM traffic parameters:
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fwan_c/wcfatm.htm#1001126

Link fragmentation and interleaving:

- MLP interleaving and queuing for Real-Time traffic (Cisco IOS Release 12.0):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/dial_c/dcppp.htm#4550
- FRF.12 (Cisco IOS Release 12.0.4T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t4/120tvofr/index.htm
- Link fragmentation and interleaving for Frame Relay and ATM virtual circuits (Cisco IOS Release 12.1.5T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t5/dtlfifra.htm
- Distributed link fragmentation and interleaving over leased lines (Cisco IOS Release 12.2.8T):
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t8/ftdlfi2.htm

- Distributed link fragmentation and interleaving for Frame Relay and ATM interfaces (Cisco IOS Release 12.2.4T):
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t4/ftdlfi.htm

Compressed Real-Time Protocol:

- RTP and TCP header compression (Cisco IOS Release 12.0.7T):
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t7/rtpfast.htm
- Class-based RTP and TCP header compression (Cisco IOS Release 12.2.13T):
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/fthdrcmp.htm

Tx-ring:

- Tx-ring tuning:
  http://www.cisco.com/en/US/tech/tk39/tk824/technologies_tech_note09186a00800fbafc.shtml

PAK_priority:

- Understanding how routing updates and Layer 2 control packets are queued on an interface with a QoS service policy:
  http://www.cisco.com/warp/public/105/rtgupdates.html

ISDN:

- Multiclass Multilink PPP (Cisco IOS Release 12.2.13T)
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftmmlppp.htm

Marking:

- Class-based marking (Cisco IOS Release 12.1.5T):
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t5/cbpmark2.htm
- Enhanced packet marking (Cisco IOS Release 12.2.13T):
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftenpkmk.htm

CHAPTER 4

Branch Router QoS Design

This chapter discusses Branch QoS considerations and designs, including the following:

- AutoQoS—Enterprise
- Unidirectional Applications
- Branch LAN edge ingress classification
- Branch NBAR policies for worm identification and policing

Chapter 3, "WAN Aggregator QoS Design," discussed the QoS design recommendations for WAN aggregators in detail. For the most part, theses 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. This chapter examines in detail these considerations and their related designs.

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

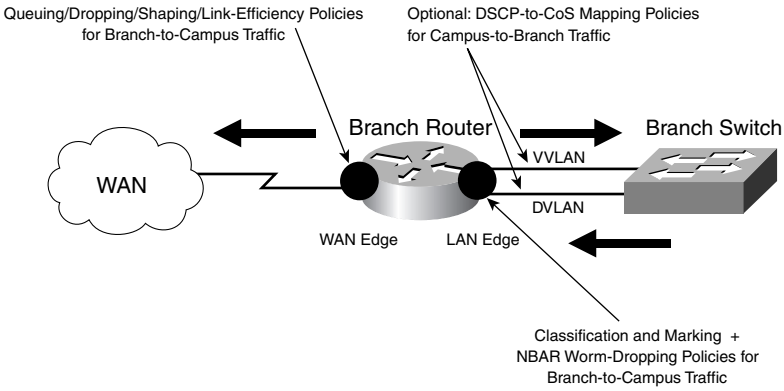
If QoS policies are to be defined manually, then other considerations must also be taken into account, such as the presence of unidirectional applications. Some applications, such as Streaming-Video (whether unicast or multicast), require bandwidth allocation only on the WAN aggregator's WAN edge, not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

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 branch router's LAN edge 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 unique consideration 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, RPC DCOM/W32/MS Blaster, Sasser, and other worms.

Figure 4-1 shows the QoS policies required on a remote branch router.

Figure 4-1 Branch Router QoS Policies



Branch WAN Edge QoS Design

WAN edge considerations discussed in Chapter 3, "WAN Aggregator QoS Design" apply also to the branch router's WAN edge. This includes the following:

- Link-speed categories (slow, medium, and high) and their respective design implications
- Bandwidth-provisioning guidelines (at least 25 percent for Best-Effort traffic and no more than 33 percent for Real-Time applications)
- Link-specific caveats (such as the fact that tools such as LFI and cRTP, when enabled, must be enabled on both ends of the WAN link for them to function correctly)

Distributed platform idiosyncrasies are not as relevant for branch router designs because these platforms rarely are deployed at remote sites (although they can be used for site-to-site links, as discussed in Chapter 3, "WAN Aggregator QoS Design").

AutoQoS—Enterprise

For small-to-medium businesses supported by administrators with moderate QoS expertise, Cisco AutoQoS—Enterprise may be an expeditious option to enable QoS for WAN/Branch networks. This advanced IOS feature detects and automatically provisions bandwidth for up to 10 classes of traffic on a link-by-link basis based on an analysis performed on sampled traffic rates. Some considerations to keep in mind when deciding whether to use the AutoQoS—Enterprise feature:

- **Policies vary on a link-by-link basis**—AutoQoS—Enterprise samples traffic rates by application classes on a link-by-link basis and then presents a configuration recommendation based on these sampled rates. These rates vary link-by-link, so the recommended policies also vary link-by-link. AutoQoS— Enterprise is not a suitable tool for enterprises that want to centralize QoS policies for consistent end-to-end service-levels for their traffic classes.
- **Policies are reactive to traffic conditions during the sampling period, rather than administratively proactive**—The sampled traffic rates, on which the AutoQoS—Enterprise feature bases its recommendations, may or may not be inline with the desired business objectives of the QoS deployment. For example, AutoQoS—Enterprise may detect and provision bandwidth for a Streaming-Video class, but perhaps Streaming-Video is not viewed by the enterprise as a critical

application requiring explicit bandwidth guarantees; they might instead prefer to use the bandwidth to increase the service-levels to Transactional Data applications. In such a case, policies would have to be manually defined.

- **AutoQoS— Enterprise does not support a Mission-Critical Data class**—If an enterprise requires the support of a Mission-Critical Data class, which is a subjective evaluation of relative business priority of Transactional/Interactive applications (as discussed in Chapter 1, "Quality of Service Design Overview"), such a class would have to be manually defined.
- **AutoQoS—Enterprise can be used as a template**—The policies proposed by AutoQoS—Enterprise can be manually modified and tweaked to meet custom requirements. The configuration generated by AutoQoS—Enterprise is not an "all or nothing" option, but rather is modifiable and thus can be viewed as a template.
- **Incremental CPU load of NBAR**—Because AutoQoS—Enterprise relies heavily on NBAR—which generates a degree of incremental CPU load—we generally do not recommended that you enable it on large WAN networks, in particular on WAN Aggregation routers serving a large number of remote branches. AutoQoS— Enterprise is therefore more suitable for smaller WAN/Branch environments.
- **Link-specific restrictions**—AutoQoS— Enterprise has several link-specific restrictions, including:
  - **Serial Interface Restrictions**—For a serial interface with a low-speed link, MLP is configured automatically. The serial interface must have an IP address. When MLP is configured, this IP address is removed and put on the MLP bundle. To ensure that the traffic goes through the low-speed link, the following conditions must be met:
    - The AutoQoS for the Enterprise feature must be configured at both ends of the link.
    - The amount of bandwidth configured must be the same on both ends of the link.
  - **Frame Relay DLCI Restrictions**:
    - This feature cannot be configured on a Frame Relay DLCI if a map class is attached to the DLCI.
    - If a Frame Relay DLCI is already assigned to one subinterface, the AutoQoS for the Enterprise feature cannot be configured for a different subinterface.
    - For low-speed Frame Relay DLCIs configured for use on Frame Relay-to-ATM networks, MLP over Frame Relay (MLPoFR) is configured automatically. The subinterface must have an IP address. When MLPoFR is configured, this IP address is removed and put on the MLP bundle. The AutoQoS for the Enterprise feature must also be configured on the ATM side of the network.
    - For low-speed Frame Relay DLCIs with Frame Relay-to-ATM Interworking, the AutoQoS for the Enterprise feature cannot be configured if a virtual template is already configured for the DLCI.
  - **ATM PVC Restrictions**:
    - For a low-speed ATM PVC, the AutoQoS for the Enterprise feature cannot be configured if a virtual template is already configured for the ATM PVC.
    - For low-speed ATM PVCs, MLP over ATM (MLPoATM) is configured automatically. The subinterface must have an IP address. When MLPoATM is configured, this IP address is removed and put on the MLP bundle. The AutoQoS for the Enterprise feature must also be configured on the ATM side of the network.

The AutoQoS for the Enterprise feature consists of two configuration phases, completed in the following order:

1. **Auto-Discovery (data collection)**—The Auto-Discovery phase can be configured to operate in one of two modes:
   - **Untrusted Mode**—In untrusted mode (configured using the **auto discovery qos** interface command), the Auto-Discovery phase uses NBAR protocol discovery to detect the applications on the network and performs statistical analysis on the network traffic.
   - **Trusted Mode**—In trusted mode (configured using the **auto discovery qos trust** interface command), the Auto-Discovery phase classifies packets based on DSCP values in the IP header and collects the statistics to calculate bandwidth and average rate/peak rate and passes that data to the template module.
2. **AutoQoS Template Generation and Installation**—This phase (configured using the **auto qos** command) generates templates from the data collected during the Auto-Discovery phase and installs the templates on the interface. These templates are then used as the basis for creating the class maps and policy maps for your network. After the class maps and policy maps are created, they are then installed on the interface.

AutoQoS— Enterprise discovers and classifies as many as 10 classes of traffic, based on the QoS Baseline model. The only traffic class not automatically detected and classified is the Mission-Critical Data class, as this class requires a subjective evaluation on relative business priority vis-à-vis other Transactional/Interactive Data applications (as described in Chapter 1, "Quality of Service Design Overview").

These AutoQoS classes are described in Table 4-1.

Table 4-1 AutoQoS for the Enterprise Feature Class Definitions

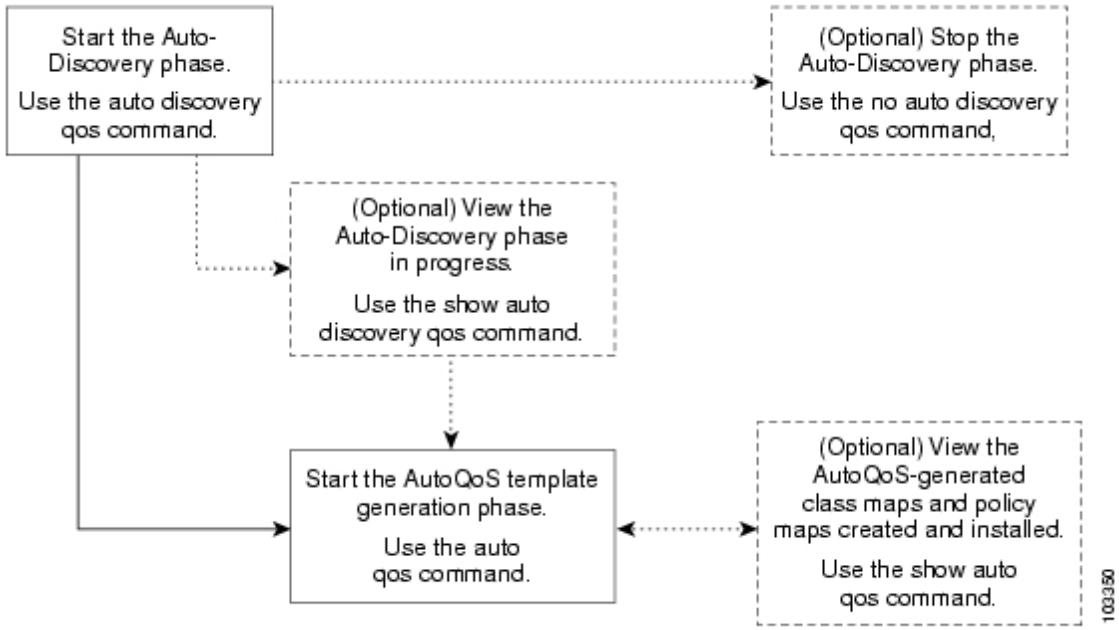
AutoQoS Class Name	Traffic Type	DSCP Value
IP Routing	Network control traffic, such as routing protocols	CS6
Interactive Voice	Inactive voice-bearer traffic	EF
Interactive Video	Interactive video data traffic	AF41
Streaming Video	Streaming media traffic	CS4
Telephony Signaling	Telephony signaling and control traffic	CS3
Transactional/Interactive	Database applications transactional in nature	AF21
Network Management	Network management traffic	CS2
Bulk Data	Bulk data transfers; web traffic; general data service	AF11
Scavenger	Casual entertainment; rogue traffic; traffic in this category is given less-than-best-effort treatment	CS1
Best Effort	Default class; all non-critical traffic; HTTP; all miscellaneous traffic	0

Additionally, the AutoQoS—Enterprise feature includes the following verification commands:

- **show auto discovery qos**
- **show auto qos**

The top-level processes for configuring and verifying the AutoQoS—Enterprise feature are illustrated in Figure 4-2.

Figure 4-2 Top-Level Processes for Configuring the AutoQoS for the Enterprise Feature



As noted, the AutoQoS— Enterprise feature is a handy alternative for automating QoS deployment in smaller WAN/Branch environments. For more information, see the IOS documentation:

http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123newft/123t/123t_11/ft_aqose.htm

If, on the other hand, all or some of the branch QoS policies must be defined manually, then additional considerations must be taken into account, as discussed below.

Unidirectional Applications

Some applications are completely symmetrical and require identical bandwidth provisioning on both ends of the WAN link. For example, if 100 kbps of LLQ are assigned to voice in one direction, 100 kbps of LLQ also must be provisioned for voice in the opposite direction (assuming that the same VoIP codecs are being used in both directions, and putting aside for a moment multicast Music-on-Hold [MoH] provisioning). Furthermore, having symmetrical polices on both sides of the WAN links greatly simplifies QoS policy deployment and management, which is an important aspect of large-scale designs.

However, certain applications, such as Streaming-Video and multicast MoH, most often are unidirectional. Therefore, it might be unnecessary and even inefficient to provision any bandwidth guarantees for such traffic on the branch router for the branch-to-campus direction of traffic flow.

Most applications lie somewhere in the middle of the scale, between the extremes of being fully bidirectional and being completely unidirectional. Most client/server applications lie closer to the unidirectional end of the scale because these applications usually consist of small amounts of client-to-server traffic coupled with larger amounts of server-to-client traffic. Such behavior can be reflected in the asymmetrical bandwidth provisioning for such types of applications.

For purely unidirectional applications, it is recommended that provisioning be removed from the WAN edge policies on branch routers and the allocated bandwidth be redistributed among other classes.

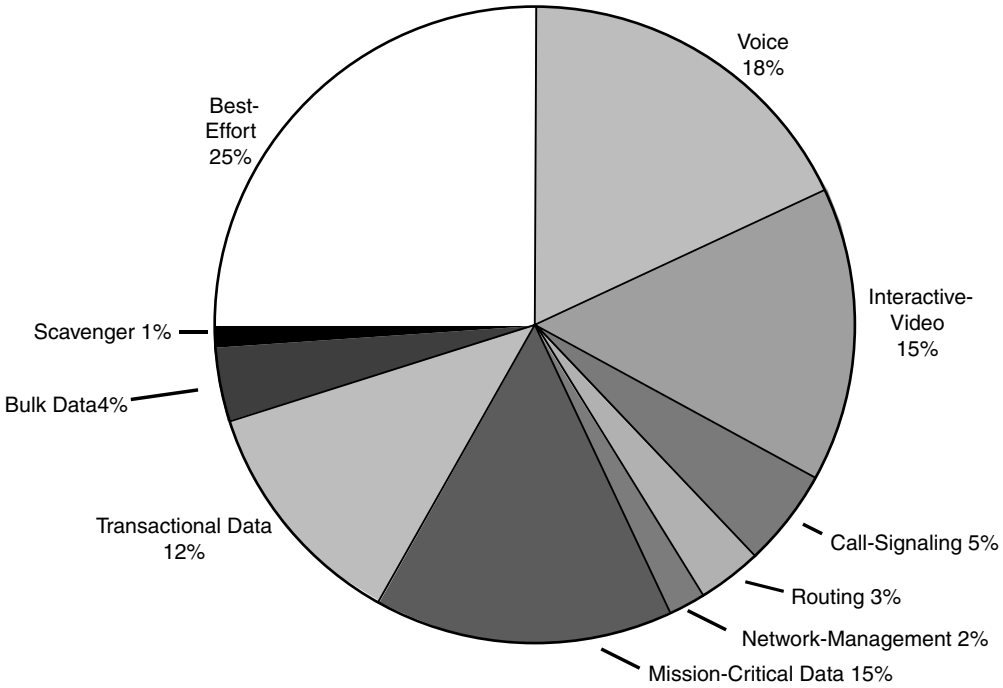
Branch Router WAN Edge (10-Class) QoS Baseline Model

The inclusion or exclusion of the Streaming-Video class affects only those enterprises deploying complex QoS class models, such as the 11-Class QoS Baseline Model. When the Streaming-Video class is removed from this model (for branch router WAN edges), the bandwidth previously allocated to this class can be reallocated among the other preferential data classes, as illustrated in Figure 4-3. Notice that no class is provisioned for Streaming-Video in this model, and the bandwidth assigned to it from the 11-Class WAN edge model (from Figure 3-6 in the previous chapter) has been reassigned to the Mission-Critical Data class and the Transactional Data class.

The configuration for such a branch router (10-class) QoS Baseline policy on a dual-T1 interface is shown in Example 4-1. Notice that there is no Streaming-Video class and that the bandwidth for it has been distributed equally between the Mission-Critical Data class (now at 15 percent instead of 10 percent) and the Transactional Data class (now at 12 percent instead of 7 percent).

A **bandwidth** statement is used on class-default, precluding the use of **fair-queue** on the class for all nondistributed platforms. Also, a **max-reserved-bandwidth 100** command must be issued on the interface before the **service-policy output** command.

Figure 4-3 Branch Router (10-Class) QoS Baseline WAN Edge Model Bandwidth Allocations Example (Dual-T1 Link Example)



Example 4-1 Branch Router (10-Class) QoS Baseline WAN Edge Model

```
!
class-map match-all VOICE
  match ip dscp ef            ! IP Phones mark Voice to EF
class-map match-all INTERACTIVE-VIDEO
  match ip dscp af41 af42     ! Recommended markings for IP/VC
class-map match-any CALL-SIGNALING
  match ip dscp cs3           ! Future Call-Signaling marking
  match ip dscp af31          ! Current Call-Signaling marking
class-map match-all ROUTING
```

```
  match ip dscp cs6           ! Routers mark Routing traffic to CS6
class-map match-all NET-MGMT
  match ip dscp cs2           ! Recommended marking for Network Management
class-map match-all MISSION-CRITICAL-DATA
  match ip dscp 25            ! Interim marking for Mission-Critical Data
class-map match-all TRANSACTIONAL-DATA
  match ip dscp af21 af22     ! Recommended markings for Transactional Data
class-map match-all BULK-DATA
  match ip dscp af11 af12     ! Recommended markings for Bulk Data
class-map match-all SCAVENGER
  match ip dscp cs1           ! Recommended marking for Scavenger traffic
!
 policy-map BRANCH-WAN-EDGE
  class VOICE
   priority percent 18        ! Voice gets 552 kbps of LLQ
  class INTERACTIVE-VIDEO
   priority percent 15        ! 384 kbps IP/VC needs 460 kbps of LLQ
  class CALL-SIGNALING
   bandwidth percent 5        ! Minimal BW guarantee for Call-Signaling
  class ROUTING
   bandwidth percent 3        ! Routing class gets 3% explicit BW guarantee
  class NET-MGMT
   bandwidth percent 2        ! Net-Mgmt class gets 2% explicit BW guarantee
  class MISSION-CRITICAL-DATA
   bandwidth percent 15       ! Mission-Critical class gets min 15% BW guarantee
   random-detect              ! Enables WRED on Mission-Critical Data class
  class TRANSACTIONAL-DATA
   bandwidth percent 12       ! Transactional-Data class gets min 12% BW guarantee
   random-detect dscp-based   ! Enables DSCP-WRED on Transactional-Data class
  class BULK-DATA
   bandwidth percent 4        ! Bulk Data class gets 4% BW guarantee
   random-detect dscp-based   ! Enables DSCP-WRED on Bulk-Data class
  class SCAVENGER
   bandwidth percent 1        ! Scavenger class is throttled
  class class-default
   bandwidth percent 25       ! Default class gets min 25% BW guarantee
   random-detect              ! Enables WRED on the default class
!
```

Verification commands:

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

Now that considerations and designs for the WAN edge of the branch router have been addressed, the next section discusses the LAN edge.

Branch Router LAN Edge QoS Design

The LAN edge of the branch router can have egress and ingress policies. Because you have been dealing with egress policies since the WAN/branch discussion began, and because the egress policies are not only optional, but also considerably simpler, they are discussed first.

As previously mentioned, it is better to mark at Layer 3 (DSCP) instead of Layer 2 whenever possible because Layer 2 markings are lost when the transmission medium changes. This is the case with any Ethernet 802.1Q/p CoS values that have been set within the campus and are carried over a WAN (or VPN). In some cases, network administrators prefer to have these markings restored at the branch; DSCP-to-CoS mapping then can be performed on the branch router's LAN edge.

In the ingress direction, the branch router might be required to perform classification and marking of branch-to-campus traffic. This might be because the branch switch lacks the capability to classify and mark traffic, or because the traffic consists of stateful flows that require NBAR for classification. NBAR classification also is required at branch LAN ingress edges to identify (and immediately drop) known worm traffic.

Each of these types of polices is discussed in more detail in the following sections.

DSCP-to-CoS Remapping

DSCP-to-CoS remapping is optional. Newer Catalyst switches perform QoS based on internal DSCP values that are generated either by trusted DSCP markings or by trusted CoS markings (coupled with CoS-to-DSCP mappings). In the case of legacy switches at the branch that perform QoS strictly by preset CoS values, CoS might need to be remapped on the branch router's LAN edge.

Enhanced packet marking (Cisco IOS Release 12.2[13]T or higher) is the optimal tool for resetting CoS values because it uses a table. In this manner, a default DSCP-to-CoS mapping can be used without having to configure explicitly a class-based marking policy that matches every DiffServ class and performs a corresponding **set cos** function.

In both cases, keep in mind that only Ethernet trunking protocols, such as 802.1Q, carry CoS information. Therefore, the policy works correctly only when applied to a trunked subinterface, not to a main Ethernet interface.

Example 14-2 presents an enhanced packet marking DSCP-to-CoS configuration for a branch LAN edge. Note that this DSCP-to-CoS remapping policy requires application to both the voice VLAN (VVLAN) subinterface and the data VLAN (DVLAN) subinterface on the branch router's LAN edge.

Example 4-2 Branch LAN Edge Enhanced Packet Marking for DSCP-to-CoS Remapping Example

```
!
ip cef                                  ! IP CEF is Required for Packet Marking
!
policy-map BRANCH-LAN-EDGE-OUT
  class class-default
   set cos dscp                         ! Enables default DSCP-to-CoS Mapping
!
!
interface FastEthernet0/0
 no ip address
 speed auto
 duplex auto
!
interface FastEthernet0/0.60
 description DVLAN SUBNET 10.1.60.0
 encapsulation dot1Q 60
 ip address 10.1.60.1 255.255.255.0
 service-policy output BRANCH-LAN-EDGE-OUT       ! Restores CoS for Data VLAN
!
interface FastEthernet0/0.160
 description VVLAN SUBNET 10.1.160.0
 encapsulation dot1Q 160
 ip address 10.1.160.1 255.255.255.0
 service-policy output BRANCH-LAN-EDGE-OUT        ! Restores CoS on Voice VLAN
!
```

Verification commands:

- **show policy**

- **show policy interface**

Branch-to-Campus Classification and Marking

In keeping with the unofficial Differentiated Services design principle of marking traffic as close to its source as possible, IP phones mark voice-bearer traffic (to DSCP EF) and Call-Signaling traffic (currently, to DSCP AF31, but this will soon change to DSCP CS3) on the phones themselves. Some IP/VC devices mark Interactive-Video traffic to AF41 on their network interface cards (NICs).

However, as has already been discussed, it is generally not recommended that end-user PCs be trusted to set their CoS/DSCP markings correctly because users easily can abuse this (either unintentionally or deliberately).

Note There may be some exceptions to this general rule, such as PCs running applications like Cisco VT Advantage in controlled environments. A network administrator must make these policy decisions.

Therefore, application traffic that originates from untrusted hosts should be marked on branch access switches. However, in some circumstances, this might not be possible:

- The branch access switch does not have Layer 3 or Layer 4 awareness for traffic classification or does not support marking.
- Classification needs to be performed at the application layer (through NBAR).

In such cases, DSCP classification must be performed at the ingress interface of the branch router.

Administrators can identify and then mark application traffic based on the following criteria:

- **Source or destination IP address (or subnet)**—Typically, destination subnets are used when setting branch QoS policies. For example, the destination subnet for a group of application servers can be used to identify a particular type of client-to-server application traffic.
- **Well-known TCP/UDP ports**—It is important to know whether the well-known port is a source port or a destination port from the branch router's perspective.
- **NBAR protocol (for example, Citrix or KaZaa) or application subparameter (for example, HTTP URL)**—NBAR Packet Description Language Modules (PDLMs) also can be configured to identify stateful or proprietary applications, as well as worms.

Regardless of the method used to identify the application traffic, the inbound classification and marking policy needs to be applied only to the DVLAN subinterface. This is because only trusted IP Telephony applications, which mark their voice traffic correctly, are admitted onto the voice VLAN.

Note If the branch access switches support access-edge policers (as described in Chapter 2, "Campus QoS Design"), these likewise should be enabled on them. This adds another layer of defense to the network, mitigating DoS/worm traffic that originates from the branch through Scavenger-class QoS.

If the branch access switches do not support such policing, compatible policers could be placed at the branch router access edge. This is in harmony with the principle of policing as close to the source as possible.

Whenever possible, though, such policing should be done in Catalyst hardware rather than Cisco IOS Software.

Although it might be tempting to use the same class-map names (such as MISSION-CRITICAL, TRANSACTIONAL-DATA, or BULK-DATA) for ingress LAN edge classification and marking policies as are in use for egress WAN edge queuing policies, this might cause confusion in policy definition and troubleshooting. Therefore, for management and troubleshooting simplicity, it is beneficial to have similar yet descriptive names for these new classes (for example, branch-originated traffic might have class-map names prepended with "BRANCH-"). A description can also be added to the class maps with the **description** command.

Source or Destination IP Address Classification

Example 4-3 shows how traffic destined to a specific subnet (10.200.200.0/24)—which, in this case, represents a server farm of proprietary Mission-Critical Data application servers—can be identified and marked on ingress on the branch router's LAN edge.

Example 4-3 Branch LAN Edge Destination IP Classification and Marking Example

```
!
ip cef                                                  ! Required for Packet Marking
!
class-map match-all BRANCH-MISSION-CRITICAL
  match access-group name MISSION-CRITICAL-SERVERS    ! ACL to reference
!
 policy-map BRANCH-LAN-EDGE-IN
  class BRANCH-MISSION-CRITICAL
   set ip dscp 25   ! (Interim) Recommended marking for Mission-Critical traffic
!
...
!
interface FastEthernet0/0
 no ip address
 speed auto
 duplex auto
!
interface FastEthernet0/0.60
 description DVLAN SUBNET 10.1.60.0
 encapsulation dot1Q 60
 ip address 10.1.60.1 255.255.255.0
 service-policy output BRANCH-LAN-EDGE-OUT  ! Restores CoS on Data VLAN
 service-policy input BRANCH-LAN-EDGE-IN    ! Marks MC Data on ingress
!
...
!
ip access-list extended MISSION-CRITICAL-SERVERS
 permit ip any 10.200.200.0 0.0.0.255    ! MC Data Server-Farm Subnet
!
```

Verification commands:

- **show policy**
- **show policy interface**
- **show ip access-list**

Verification Command: show ip access-list

When access lists are used as the filtering criteria for a class map, the **show ip access-list** command is helpful in identifying whether the access list is registering matches, especially for ACLs that have multiple lines of match criteria. This command provides granular visibility into which lines of an access list are registering matches. In Example 4-4, 464 matches are registered against the access list.

Example 4-4 show ip access-list Verification of Remote Branch LAN Edge Destination IP Classification Example

```
BRANCH#60-C3745#show ip access-list MISSION-CRITICAL-SERVERS
Extended IP access list MISSION-CRITICAL-SERVERS
    10 permit ip any 10.200.200.0 0.0.0.255 (464 matches)
BRANCH#60-C3745#
```

Well-Known TCP/UDP Port Classification

Most applications can be identified by their well-known TCP/UDP ports. Some of these ports have keywords within Cisco IOS Software to identify them when defining access lists.

Note The Internet Assigned Numbers Authority (IANA) lists registered well-known and registered application ports at http://www.iana.org/assignments/port-numbers.

Building on Example 4-4, Example 4-5 classifies and marks branch-originated FTP and e-mail traffic, both POP3 and IMAP (TCP port 143), as Bulk Data (DSCP AF11).

Example 4-5 Branch LAN Edge Well-Known Port Classification Example

```
!
ip cef                                          ! Required for Packet Marking
!
class-map match-all BRANCH-MISSION-CRITICAL
  match access-group name MISSION-CRITICAL-SERVERS
class-map match-all BRANCH-BULK-DATA
  match access-group name BULK-DATA-APPS      ! ACL to reference
!
 policy-map BRANCH-LAN-EDGE-IN
  class BRANCH-MISSION-CRITICAL
   set ip dscp 25
  class BRANCH-BULK-DATA
   set ip dscp af11                           ! Bulk data apps are marked to AF11
!
!
interface FastEthernet0/0
 no ip address
 speed auto
 duplex auto
!
interface FastEthernet0/0.60
 description DVLAN SUBNET 10.1.60.0
 encapsulation dot1Q 60
 ip address 10.1.60.1 255.255.255.0
 service-policy output BRANCH-LAN-EDGE-OUT      ! Restores CoS on Data VLAN
 service-policy input BRANCH-LAN-EDGE-IN        ! Marks Data on ingress
!
...
```

```
!
ip access-list extended MISSION-CRITICAL-SERVERS
 permit ip any 10.200.200.0 0.0.0.255
!
ip access-list extended BULK-DATA-APPS
 permit tcp any any eq ftp                    ! Identifies FTP Control traffic
 permit tcp any any eq ftp-data               ! Identifies FTP Data traffic
 permit tcp any any eq pop3                   ! Identifies POP3 E-mail traffic
 permit tcp any any eq 143                    ! Identifies IMAP E-mail traffic
!
```

Verification commands:

- **show policy**
- **show policy interface**
- **show ip access-list**

NBAR Application Classification

At the time of this writing, Cisco IOS Software included NBAR PDLMs for 98 of the most common network applications, with the capability to define an additional 10 applications using custom PDLMs.

Of these protocols, 15 require stateful packet inspection for positive identification. Because NBAR operates in the IP Cisco Express Forwarding (CEF) switching path, only the first packet within a flow requires stateful packet inspection, and the policy is applied to all packets belonging to the flow. NBAR stateful packet inspection requires more CPU processing power than simple access control lists (ACLs). However, on newer branch router platforms, such as Cisco's 1800, 2800, and 3800 series Integrated Services Routers (ISRs), Cisco Technical Marketing testing has shown the overhead of enabling NBAR classification at dual-T1 rates to be quite minimal (typically 2 to 5 percent, depending on the traffic mix).

Building again on the previous example, Example 4-6 uses NBAR to identify several types of Transactional Data applications, Network-Management applications, and Scavenger applications.

In Example 4-6, Transactional Data applications include Citrix, LDAP, Oracle SQL*NET, and HTTP web traffic with "SalesReport" in the URL. In addition, a custom PDLM is defined to identify SAP traffic (by TCP ports 3200 through 3203 and 3600), and SAP also is included as a Transactional Data application.

Additionally, Network-Management applications, such as SNMP, Syslog, Telnet, NFS, DNS, ICMP, and TFTP, are identified through NBAR and marked to DSCP CS2.

Furthermore, Scavenger applications, such as Napster, Gnutella, KaZaa (versions 1 and 2), Morpheus, Grokster, and many other peer-to-peer file-sharing applications, are identified by NBAR and marked to DSCP CS1.

Example 4-6 Branch LAN Edge NBAR Classification Example

```
!
ip nbar port-map custom-01 tcp 3200 3201 3202 3203 3600   ! PDLM Mapping for SAP
!
ip cef     ! IP CEF is required for both Class-Based Marking and for NBAR
!
class-map match-all BRANCH-MISSION-CRITICAL
  match access-group name MISSION-CRITICAL-SERVERS
class-map match-any BRANCH-TRANSACTIONAL-DATA! Must use "match-any"
  match protocol citrix                      ! Identifies Citrix traffic
  match protocol ldap                       ! Identifies LDAP traffic
  match protocol sqlnet                     ! Identifies Oracle SQL*NET traffic
  match protocol http url "*SalesReport*"   ! Identifies "SalesReport" URLs
```

```
  match protocol custom-01        ! Identifies SAP traffic via Custom-01 PDLM Port-Map
class-map match-all BRANCH-BULK-DATA
  match access-group name BULK-DATA-APPS
class-map match-any BRANCH-NET-MGMT
  match protocol snmp                          ! Identifies SNMP traffic
  match protocol syslog                        ! Identifies Syslog traffic
  match protocol telnet                        ! Identifies Telnet traffic
  match protocol nfs                           ! Identifies NFS traffic
  match protocol dns                           ! Identifies DNS traffic
  match protocol icmp                          ! Identifies ICMP traffic
  match protocol tftp                          ! Identifies TFTP traffic
class-map match-any BRANCH-SCAVENGER
  match protocol napster                       ! Identifies Napster traffic
  match protocol gnutella                      ! Identifies Gnutella traffic
  match protocol fasttrack                     ! Identifies KaZaa (v1) traffic
  match protocol kazaa2                        ! Identifies KaZaa (v2) traffic
!
policy-map BRANCH-LAN-EDGE-IN
  class BRANCH-MISSION-CRITICAL
   set ip dscp 25
  class BRANCH-TRANSACTIONAL-DATA
   set ip dscp af21           ! Transactional Data apps are marked to DSCP AF21
  class BRANCH-NET-MGMT
   set ip dscp cs2            ! Network Management apps are marked to DSCP CS2
  class BRANCH-BULK-DATA
   set ip dscp af11
  class BRANCH-SCAVENGER
   set ip dscp cs1             ! Scavenger apps are marked to DSCP CS1
!
...
!
interface FastEthernet0/0
 no ip address
 speed auto
 duplex auto
!
interface FastEthernet0/0.60
 description DVLAN SUBNET 10.1.60.0
 encapsulation dot1Q 60
 ip address 10.1.60.1 255.255.255.0
 service-policy output BRANCH-LAN-EDGE-OUT  ! Restores CoS on Data VLAN
 service-policy input BRANCH-LAN-EDGE-IN    ! Input Marking policy on DVLAN only
!
...
!
ip access-list extended MISSION-CRITICAL-SERVERS
 permit ip any 10.200.200.0 0.0.0.255
!
!
ip access-list extended BULK-DATA-APPS
 permit tcp any any eq ftp
 permit tcp any any eq ftp-data
 permit tcp any any eq pop3
 permit tcp any any eq 143
!
```

Note The NBAR **fasttrack** PDLM identifies KaZaa (version 1), Morpheus, Grokster, and other applications. The NBAR **gnutella** PDLM identifies Gnutella, BearShare, LimeWire, and other peer-to-peer applications.

Verification commands:

- **show policy**
- **show policy interface**
- **show ip access-list**
- **show ip nbar port-map**

Verification Command: show ip nbar port-map

When NBAR custom PDLMs are used to identify applications, it is useful to verify that the ports bound to the PDLM have been entered correctly. This information is obtained with the **show ip nbar port-map** command. In this (filtered) example, Example 4-7, the TCP ports given for SAP (3200 through 3203 and 3600) were bound correctly to the Custom-01 NBAR PDLM.

Example 4-7 **show ip nbar port-map** ***Verification Custom-PDLM TCP/UDP Port-Mapping Example***

```
BRANCH#60-C3745#show ip nbar port-map | include custom-01
port-map custom-01            tcp 3200 3201 3202 3203 3600
BRANCH#60-C3745#
```

NBAR Known-Worm Classification and Policing

Worms are nothing new. They've been around almost as long as the Internet itself (one of the first Internet worms was the Morris worm, released in November 1988). 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.

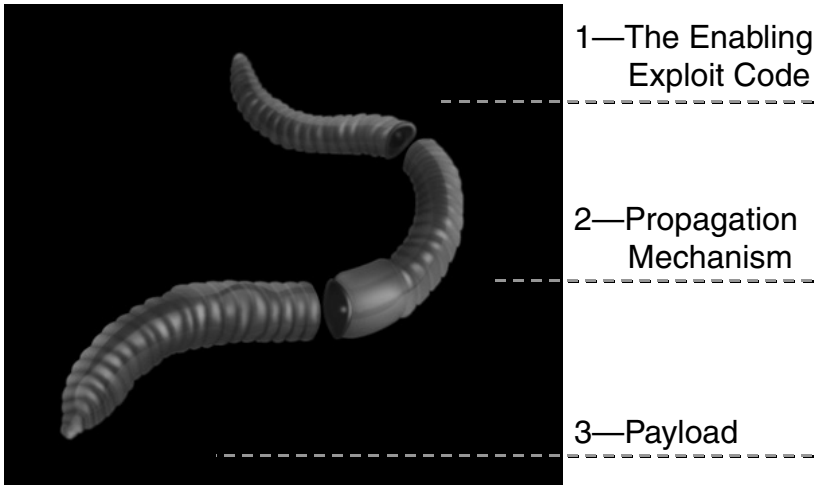
Note A virus, which is slightly different from a worm, requires a vector to carry the virus code from one system to another. The vector can be either a word-processing document, an e-mail, or an executable program.

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 (as illustrated in Figure 4-4):

- **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.

Figure 4-4 Anatomy of an Internet Worm



Some worms use unique TCP/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 branch router's LAN edge.

Other worms hijack legitimate TCP/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. Some of these known worms include Code Red, NIMDA, SQL Slammer, RPC DCOM/W32/MS Blaster, and Sasser. The following sections discuss how NBAR can be used for each of these types of 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. CodeRedv2's success 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 original Code Red payload is shown in Example 4-8. The initial infection attempt sends a large HTTP GET request to the target IIS server.

Example 4-8 Original Code Red Payload

```
2001-08-04 16:32:23 24.101.17.216 - 10.1.1.75 80 GET /default.ida
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNN%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd
```

```
3%u7801%u9090%u6858%ucbd3%u7801%u9090%u9090%u8190%u00c3%u0003%u8b00%u
531b%u53ff%u0078%u0000%u00=a 403
```

The CodeRedv2 payload is shown in Example 4-9.

Example 4-9 CodeRedv2 Payload

```
2001-08-04 15:57:35 64.7.35.92 - 10.1.1.75 80 GET /default.ida
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX%u9090%u6858%ucbd3%u7801%u
9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u9090%u
8190%u00c3%u0003%u8b00%u531b%u53ff%u0078%u0000%u00=a 403 -
```

Notice that the GET request in both cases is looking for a file named default.ida. However, this filename changes in newer variants of Code Red, such as CodeRedv3/CodeRed.C, as shown in Example 4-10.

Example 4-10 CodeRedv3 Payload

```
2001-08-06 22:24:02 24.30.203.202 - 10.1.1.9 80 GET /x.ida
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA=X 403 HTTP/1.1 -
```

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 these criteria (.ida, cmd.exe, and root.exe), as shown in Example 4-11.

Example 4-11 NBAR Policies to Identify Code Red

```
!
class-map match-any CODE-RED
   match protocol http url "*.ida*"          ! Identifies HTTP GET .ida requests
   match protocol http url "*cmd.exe*"       ! Identifies HTTP with cmd.exe
   match protocol http url "*root.exe*"      ! Identifies HTTP with root.exe
!
```

Verification command:

- show policy

NBAR Versus NIMDA

Two months after Code Red struck the Internet, another large-scale worm, NIMDA, was released. Unlike Code Red, NIMDA was a hybrid worm because it contained the characteristics of both a worm and a virus. NIMDA spread using several vectors:

- Through e-mail as an attachment (virus vector)
- Through network shares (worm vector)
- Through JavaScript by browsing compromised websites (virus vector)

- Through infected hosts actively scanning for additional exploitable hosts (worm vector)
- Through infected hosts actively scanning for back doors created by Code Red (worm vector)

NIMDA did not appear to exhibit intentional destructive capabilities; to date, NIMDA's activities have been restricted to its self-propagation, which has the side effect of a DoS (flooding) attack.

NIMDA propagates itself by copying, downloading, or executing a file called readme.eml. Therefore, NBAR can be used to check the payload of HTTP packets to see if they are propagating this file, as shown in Example 4-12.

Example 4-12 NBAR Policies to Identify NIMDA

```
!
class-map match-any NIMDA
   match protocol http url "*readme.eml*"     ! Identifies HTTP with "readme.eml"
!
```

Verification command:

- **show policy**

NBAR Versus SQL Slammer

After the NIMDA infection subsided, the Internet saw the appearance of smaller infectious worms. 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. Example 4-13 shows a signature string from the SQL Slammer worm.

Example 4-13 SQL Slammer Worm Signature String

```
\x04\x01\x01\x01\x01\x01.*[.][Dd][Ll][Ll]
```

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 (necessary with TCP-based worms) 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. SQL Slammer's doubling rate was approximately 8.5 seconds. In contrast, CodeRedv2's doubling rate 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 Example 4-14.

Example 4-14 NBAR Policies to Identify SQL Slammer

```
!
ip nbar port-map custom-02 udp 1434          ! Maps a custom PDLM to UDP 1434
!
class-map match-all SQL-SLAMMER
   match protocol custom-02                  ! Matches the custom Slammer PDLM
   match packet length min 404 max 404       ! Matches the packet length (376+28)
!
```

Verification commands:

- show policy
- show ip nbar port-map

Note Because NBAR custom-01 PDLM has been used in previous examples to identify SAP traffic, another custom PDLM (custom-02) is used in this example. Subsequent examples similarly are defined with the next-available custom PDLM.

NBAR Versus RPC DCOM/W32/MS Blaster

First released in August 2003, the Remote Procedure Call (RPC) Distributed Component Object Model (DCOM) worm exploited a flaw in a section of Microsoft's RPC code dealing with message exchange over TCP/IP, resulting in the incorrect handling of malformed messages. This flaw was a stack-based buffer overflow occurring in a low-level DCOM interface within the RPC process listening on TCP ports 135, 139, and 445.

The DCOM protocol enables Microsoft software components to communicate with one another. This is a core function of the Windows kernel and cannot be disabled. The vulnerability results because the Windows RPC service does not properly check message inputs under certain circumstances. By sending a malformed RPC message, an attacker can cause the RPC service on a device to fail in such a way that arbitrary code could be executed. The typical exploit for this vulnerability launches a reverse-telnet back to the attacker's host to gain complete access to the target.

Successful exploitation of this vulnerability enables an attacker to run code with local system privileges. This enables an attacker to install programs; view, change, or delete data; and create new accounts with full privileges. Because RPC is active by default on all versions of the Windows operating system, any user who can deliver a malformed TCP request to an RPC interface of a vulnerable computer could attempt to exploit the vulnerability. It is even possible to trigger this vulnerability through other means, such as logging into an affected system and exploiting the vulnerable component locally.

A variant of the RPC DCOM worm is termed W32 Blaster or MS Blaster. When MS Blaster successfully exploits a host, it attempts to upload a copy of the worm program to the newly exploited host. MS Blaster uses TFTP to copy the worm program from the attacking host to the target system. MS Blaster also starts up a cmd.exe process and binds it to TCP port 4444 of the newly exploited system. This provides any attacker with direct command-line access at the local system privilege level, as discussed previously.

To access the system, the attacker needs only to telnet to TCP port 4444 on the exploited host. If the worm is successful in copying the MS Blaster program to the target, the worm exploit code modifies the system registry to ensure that the worm is restarted if the system reboots. It then launches the worm program on the newly exploited host to begin the cycle again, starting with scanning for more exploitable hosts. MS Blaster also contained code for a DoS attack. This particular attack was targeted at www.windowsupdate.com.

NBAR can be used to combat the RPC DCOM/W32/MS Blaster worm by identifying communications on TCP/UDP ports 135, 139, and 445.

By default, the NBAR **exchange** PDLM is mapped to TCP port 135; therefore, this PDLM can be used as part of the MS Blaster worm policy definition. Similarly, the NBAR **netbios** PDLM is bound by default to TCP/UDP port 139 (in addition to TCP port 137 and UDP ports 137 and 138), so this PDLM also can be used within the policy definition; specifically, the **netbios** PDLM can have its port mapping expanded to include TCP port 445 and UDP ports 135, 139, and 445, as shown in Example 4-15.

Note Alternatively, a custom PDLM can be defined for these ports (TCP/UDP 135, 139, and 445), but before this could be done, you would have to map the **exchange** and **netbios** PDLMs ports away from their defaults, to avoid conflicting PDLM port mappings.

Example 4-15 NBAR Policies to Identify RPC DCOM/W32/MS Blaster

```
!
ip nbar port-map netbios tcp 137 139 445            ! Matches TCP 137/139/445
ip nbar port-map netbios udp 135 137 138 139 445    ! Matches UDP 135/137-139/445
!
class-map match-any MS-BLASTER
   match protocol exchange                    ! Matches TCP port 135
   match protocol netbios                     ! Matches MS Blaster NetBIOS PDLM
!
```

Verification commands:

- **show policy**
- **show ip nbar port-map**

NBAR Versus Sasser

The next major worm after MS Blaster was the Sasser worm (and variants Sasser.A/B/C/D), which was released in late April 2004. Sasser exploits a flaw in the Windows Local Security Authority Service Server (LSASS) that can cause systems to crash and continually reboot, or allow a remote attacker to execute arbitrary code with local system privileges.

Sasser is very efficient in scanning: It can scan 1024 separate IP addresses simultaneously (on TCP port 445). When scanning reveals a vulnerable system, the worm exploits the LSASS vulnerability and creates a remote shell (RSH) session on TCP port 9996 back to the infecting system. Then Sasser starts an FTP server on TCP port 5554 to retrieve a copy of the worm.

Sasser can be identified through a custom NBAR PDLM listening for communication on TCP ports 445, 5554, and 9996, as shown in Example 4-16.

Note If TCP port 445 already has been bound to the **netbios** NBAR PDLM (as recommended previously in the MS Blaster worm definition), it is not necessary to include this port in the Sasser custom PDLM port mapping (because it will cause a conflict).

Example 4-16 NBAR Policies to Identify Sasser

```
!
ip nbar port-map custom-03 tcp 445 5554 9996  ! Matches on TCP 445/5554/9996
!
class-map match-all SASSER
   match protocol custom-03                   ! Matches Sasser custom PDLM
!
```

Verification commands:

- show policy
- show ip nbar port-map

NBAR Versus Future Worms

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.

A new NBAR feature (introduced in Cisco IOS Release 12.3[4]T) enables network administrators 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 administrators to define the strings that they want to search for in the application payload (for any application, not just HTTP URLs) to identify a given application.

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.

For example, consider the example of 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. Moonbeam's (fictitious) payload is shown in Example 4-17.

Example 4-17 Moonbeam Worm Payload

```
%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 Example 4-18.

Example 4-18 NBAR Policies to Identify Moonbeam

```
!
ip nbar custom MOONBEAM 8 ascii Moonbeam tcp range 21000 21999  ! "Moonbeam" PDLM
!
class-map match-all MOONBEAM-WORM
 match protocol MOONBEAM        ! Matches the "Moonbeam" custom PDLM
!
```

Verification commands:

- **show policy**
- **show ip nbar port-map**

Policing Known Worms

These are just a few examples of known worms that can be identified using NBAR. After traffic generated by known worms has been positively identified, it should not be re-marked or limited; rather, it should be dropped immediately. This can be done on ingress on branch LAN edges, as shown in Example 4-19, which combines the policies for Code Red, NIMDA, SQL Slammer, RPC DCOM/W32/MS Blaster, and Sasser. Additionally, the fictitious worm Moonbeam has been included in the policy.

Note A recursive classification is required in this policy for SQL Slammer. This is because SQL Slammer requires a **match-all** criteria for its initial classification, but for policy-management purposes, it is desired that this initial classification of SQL Slammer be lumped under a single policy (with a **match-any** criteria) to identify and drop all known worms.

Example 4-19 NBAR Branch LAN Edge Ingress Policy for Known Worms

```
!
ip nbar port-map custom-02 udp 1434           ! SQL Slammer custom PDLM
ip nbar port-map custom-03 tcp 5554 9996      ! Sasser custom PDLM
ip nbar port-map netbios tcp 137 139 445      ! MS Blaster TCP 137/139/445
ip nbar port-map netbios udp 135 137 138 139 445 ! MS Blaster UDP 135/137-139/445
ip nbar custom MOONBEAM 8 ascii Moonbeam tcp range 21000 21999  ! "Moonbeam" PDLM
!
class-map match-all SQL-SLAMMER
   match protocol custom-02                  ! Matches the SQL Slammer PDLM
   match packet length min 404 max 404       ! Matches the packet length (376+28)
!
class-map match-any WORMS
   match protocol http url "*.ida*"          ! CodeRed
   match protocol http url "*cmd.exe*"       ! CodeRed
   match protocol http url "*root.exe*"      ! CodeRed
   match protocol http url "*readme.eml*"    ! NIMDA
   match class-map SQL-SLAMMER               ! SQL Slammer class-map
   match protocol exchange                   ! MS Blaster (TCP 135)
   match protocol netbios                    ! MS Blaster NetBIOS PDLM
   match protocol custom-03                  ! Sasser custom PDLM
   match protocol MOONBEAM                   ! "Moonbeam" PDLM
!
policy-map WORM-DROP
 class WORMS
   drop                                       ! Drops all known worms
!
...
!
interface FastEthernet0/0
 no ip address
 speed auto
 duplex auto
!
interface FastEthernet0/0.60
 description DVLAN SUBNET 10.1.60.0
 encapsulation dot1Q 60
 ip address 10.1.60.1 255.255.255.0
 service-policy input WORM-DROP           ! Drops known worms (DVLAN only)
!
```

Verification commands:

- **show policy map**
- **show policy-map interface**
- **show ip nbar port-map**

Note For a case study example of Branch Router QoS design, refer to Figure 14-4 and Example 14-20 of the Cisco Press book, *End-to-End QoS Network Design* by Tim Szigeti and Christina Hattingh.

Summary

Although the QoS design recommendations for branch routers are very similar to and are related to the QoS recommendations for WAN aggregators, this chapter examined four unique considerations of branch routers.

The first consideration was whether to define policies manually or automatically, via the AutoQoS—Enterprise feature. AutoQoS— Enterprise can automatically detect and provision bandwidth for up to 10 classes of traffic and is well suited to smaller WAN/Branch networks managed by administrators with moderate QoS expertise. However for larger WAN/Branch networks, where centralized policies are generally preferred, this feature may not be appropriate due to its limitations.

The second consideration is whether applications provisioned over the WAN are bidirectional or unidirectional. Some unidirectional applications, such as Streaming-Video, provisioned on the WAN aggregator WAN edge do not need to be provisioned correspondingly on the branch router's WAN edge. Bandwidth from unidirectional application classes can be redistributed among other preferential application classes on the branch router's WAN edge.

The third consideration unique to branch routers is that ingress marking might need to be performed on branch-to-campus traffic. This might be because the remote branch access switch does not have the capability to classify and mark traffic, or it might be because some applications require stateful packet inspection (NBAR) to identify them correctly. In either case, ingress marking policies would be required on the branch router's LAN edge, on the data VLAN's subinterface (the voice VLAN traffic markings are trusted). Branch-to-campus traffic can be identified by Layer 3 parameters (such as destination subnet), Layer 4 parameters (such as well-known TCP/UDP ports), or NBAR PDLMs. An example of each type of classification is provided. Additionally, optional DSCP-to-CoS mapping policies (for restoring 802.1p CoS markings that were lost when campus-originated traffic traversed the WAN media) can be set on the branch router's LAN edge.

A fourth unique consideration in branch QoS design is that branch router ingress LAN edges 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 worms. This chapter discussed an extension of the NBAR feature that enables administrators to program the strings that they want NBAR to search packet payloads for; this feature enables NBAR policies to be used to identify new worms that undoubtedly will be released in the future.

References

Standards

- RFC 2474 "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers" http://www.ietf.org/rfc/rfc2474
- RFC 2597 "Assured Forwarding PHB Group" http://www.ietf.org/rfc/rfc2597
- RFC 3246 "An Expedited Forwarding PHB (Per-Hop Behavior)" http://www.ietf.org/rfc/rfc3246

Books

- Szigeti, Tim and Christina Hattingh. *End-to-End QoS Network Design: Quality of Service in LANs, WANs and VPNs*. Indianapolis: Cisco Press, 2004.

Cisco IOS Documentation

- Class-based marking (Cisco IOS Release 12.1.5T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t5/cbpmark2.htm
- Enhanced packet marking (Cisco IOS Release 12.2.13T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftenpkmk.htm
- NBAR overview (Cisco IOS Release 12.3): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt1/qcfclass.htm#1003102
- Network-Based Application Recognition (Cisco IOS Release 12.1.5T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt1/qcfnbar.htm
- Network-Based Application Recognition (Cisco IOS Release 12.2.8T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t8/dtnbarad.htm
- Network-Based Application Recognition (Cisco IOS Release 12.3[4]T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t8/dtnbarad.htm
- AutoQoS for the Enterprise (Cisco IOS Release 12.3[11]T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123newft/123t/123t_11/ft_aqose.htm

Cisco SAFE Whitepapers

- SAFE worm mitigation: http://www.cisco.com/en/US/netsol/ns340/ns394/ns171/ns128/networking_solutions_white_paper09186a00801e120c.shtml
- Using Network-Based Application Recognition and ACLs for blocking the Code Red worm: http://www.cisco.com/en/US/products/hw/routers/ps359/products_tech_note09186a00800fc176.shtml
- How to protect your network against the NIMDA virus: http://www.cisco.com/en/US/products/sw/iosswrel/ps1835/products_tech_note09186a0080110d17.shtml
- SAFE SQL Slammer worm attack mitigation: http://www.cisco.com/en/US/netsol/ns340/ns394/ns171/ns128/networking_solutions_white_paper09186a00801cd7f5.shtml
- SAFE RPC DCOM/W32/Blaster attack mitigation: http://www.cisco.com/en/US/netsol/ns340/ns394/ns171/ns128/networking_solutions_white_paper09186a00801b2391.shtml
- Combating the Internet worm Sasser: http://www.cisco.com/application/pdf/en/us/guest/netsol/ns441/c664/cdccont_0900aecd800f613b.pdf

CHAPTER 5

MPLS VPN QoS Design

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

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

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

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

The following topics are discussed in this design chapter:

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

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

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

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

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

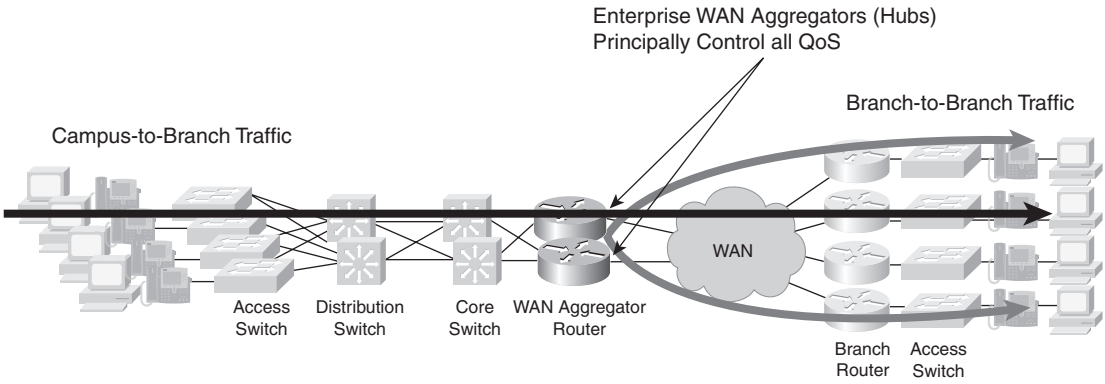
Note This chapter addresses QoS design for MPLS VPNs, not the theory and operation of MPLS VPNs themselves. It is assumed that the reader is familiar with basic MPLS VPN architectures and technologies. For a detailed discussion of MPLS VPNs, refer to the Cisco Press books *MPLS and VPN Architectures*,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

Where Is QoS Needed over an MPLS VPN?

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

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

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



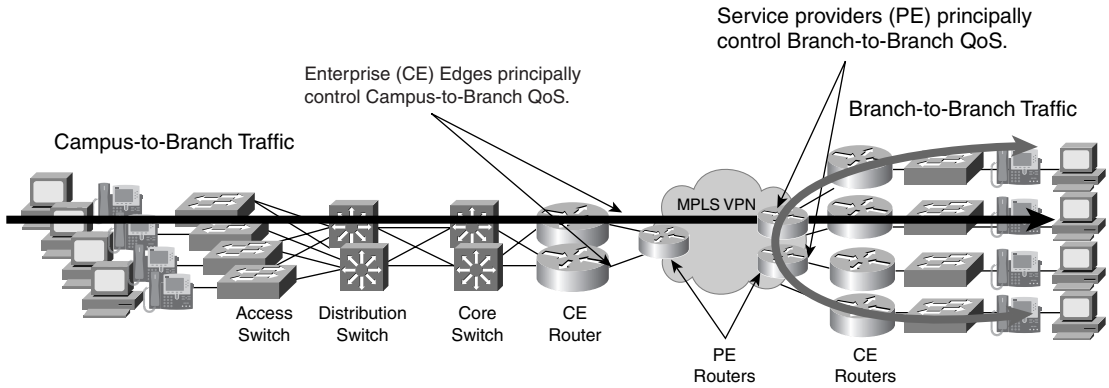
However, with the advent of MPLS VPN service offerings that inherently offer full-mesh connectivity, the QoS administration paradigm shifts. Under a full-mesh design, the hub router still administers QoS for all campus-to-branch traffic, but it no longer fully controls the QoS for branch-to-branch traffic.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

For example, consider the case of provisioning any-to-any videoconferencing. As with a traditional Layer 2 WAN design, a scheduling policy to prioritize IP/VC on the WAN aggregator is required. Then the enterprise must properly provision similar priority scheduling for IP/VC on the branch routers also. In this manner, any videoconferencing calls from the campus to the branch (and also from branch to branch) are protected against traffic of lesser importance flowing between the *same* sites. The complexity of the fully meshed model arises when considering that contending traffic might not always come for the same sites, but could come from *any* site. Furthermore, the enterprise no longer fully

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

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

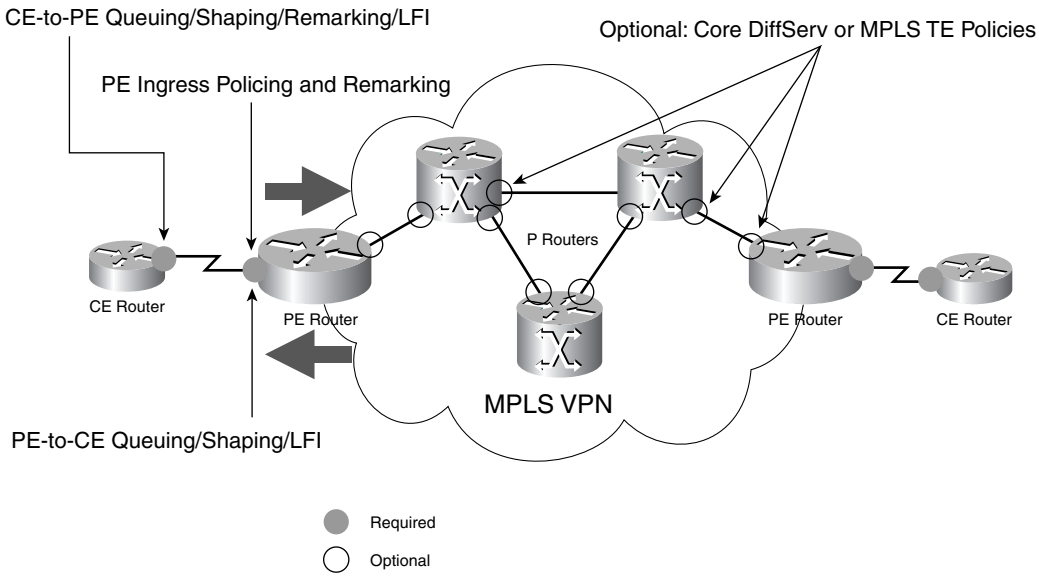
Figure 5-2 QoS Administration in Fully Meshed MPLS VPN Design



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

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

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



This design chapter focuses on CE and PE QoS design.

Customer Edge QoS Design Considerations

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

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

The following sections examine these considerations in more detail.

Layer 2 Access (Link-Specific) QoS Design

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

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

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

Service Provider Service-Level Agreements

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

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

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

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

Figure 5-4 illustrates the interrelationship of these SLAs.

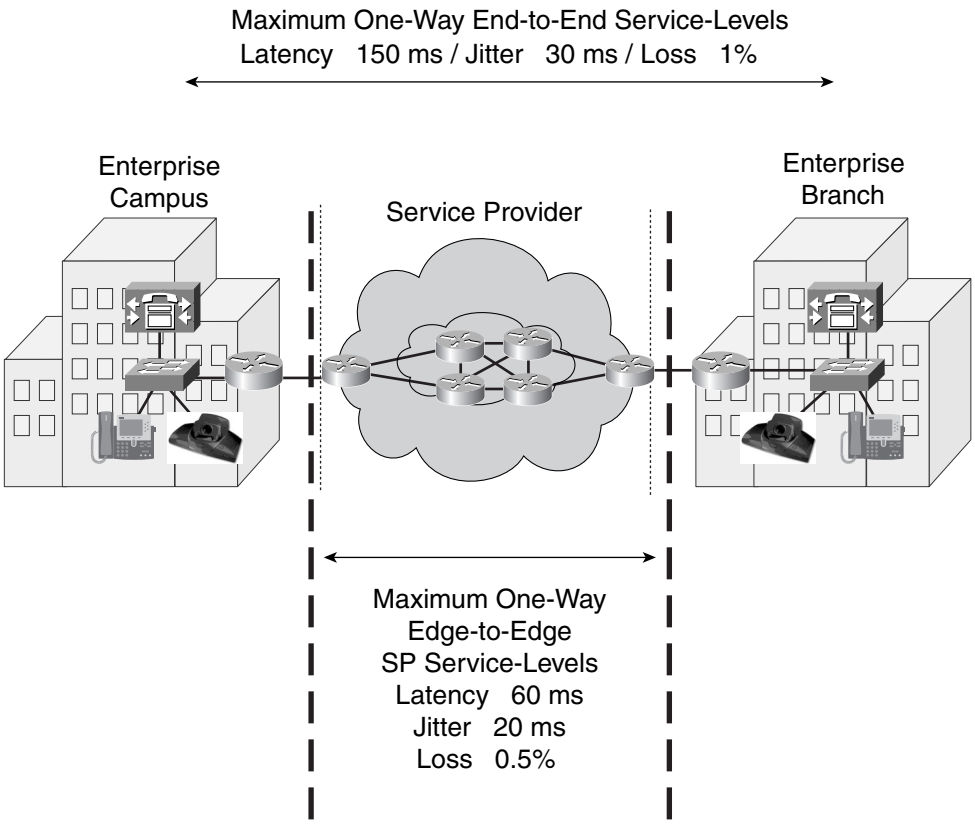
CPN-IP Multiservice Service Providers that meet these SLAs can be found at

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

Choose the IP VPN Multiservice option.

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

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



Enterprise-to-Service Provider Mapping Models

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

Voice and Video

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

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

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

Call-Signaling

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

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

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

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

Mixing TCP with UDP

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

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

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

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

Marking and Re-Marking

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

A general DiffServ principle is to mark or trust traffic as close to the source as administratively and technically possible. However, certain traffic types might need to be re-marked before handoff to the service provider to gain admission to the correct class. If such re-marking is required, it is recommended

that the re-marking be performed at the CE's egress edge, not within the campus. This is because service-provider service offerings likely will evolve or expand over time, and adjusting to such changes will be easier to manage if re-marking is performed only at the CE egress edge.

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

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

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

Verification commands:

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

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

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

Continuing and expanding on the previous example, the enterprise wants to restore the original markings that it set for Interactive-Video and Call-Signaling. Additionally, it wants to restore original markings for Oracle traffic (which it originally marked DSCP 25 and is using TCP port 9000 with) and DLSw+ traffic (originally marked AF21). Both of these data applications were handed off to the service provider marked as AF21, but they might have been marked down to AF22 within the service-provider cloud. Example 5-2 shows a configuration that enables such re-marking from the MPLS VPN. The "match-all"

criteria of the class maps performs a logical AND operation against the potential markings and re-markings, and the access list (or NBAR-supported protocol) that sifts the applications apart. The policy is applied on the same CE link, but in the ingress direction.

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

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

Verification commands:

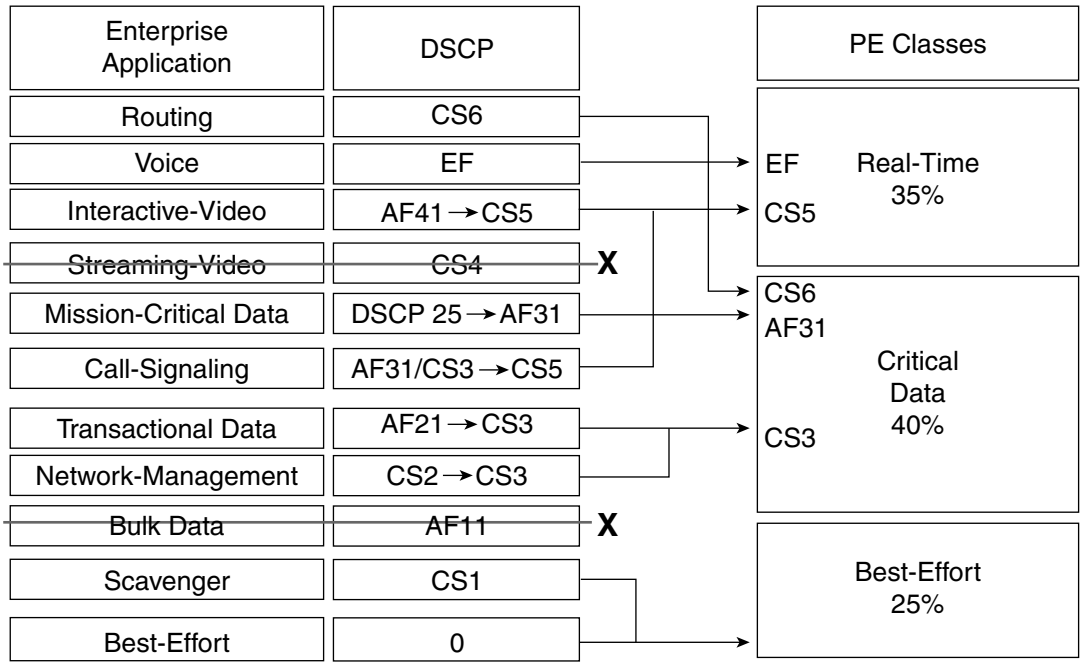
- **show policy**
- **show policy interface**

Three-Class Provider-Edge Model: CE Design

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

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

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



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

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

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

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

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

Verification commands:

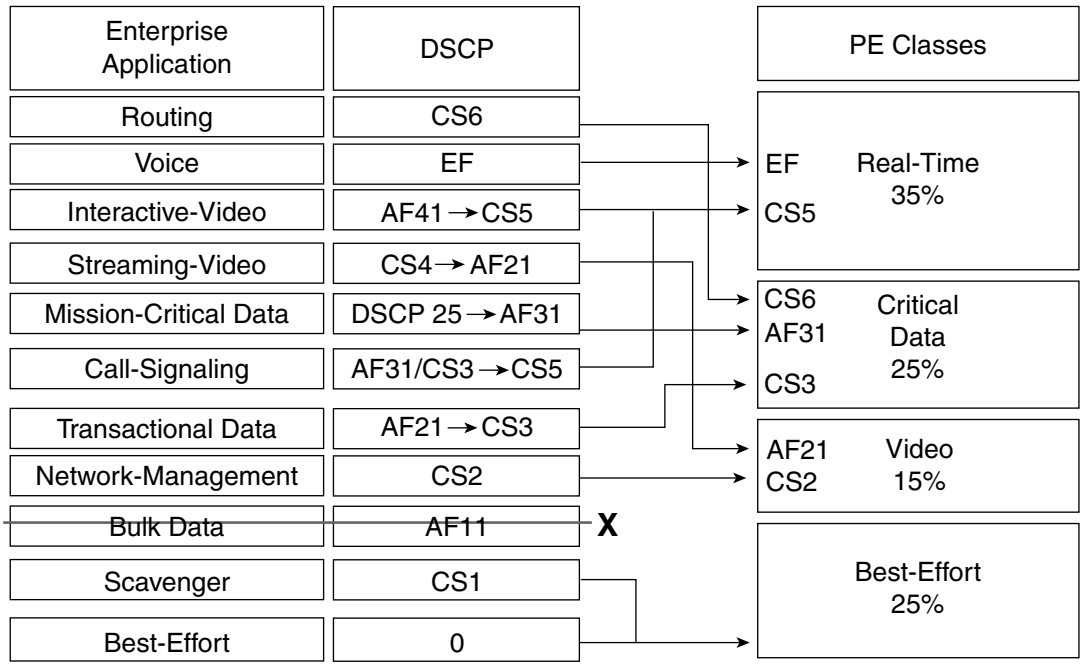
- **show policy**
- **show policy interface**

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

Four-Class Provider-Edge Model: CE Design

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

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



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

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

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

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

Verification commands:

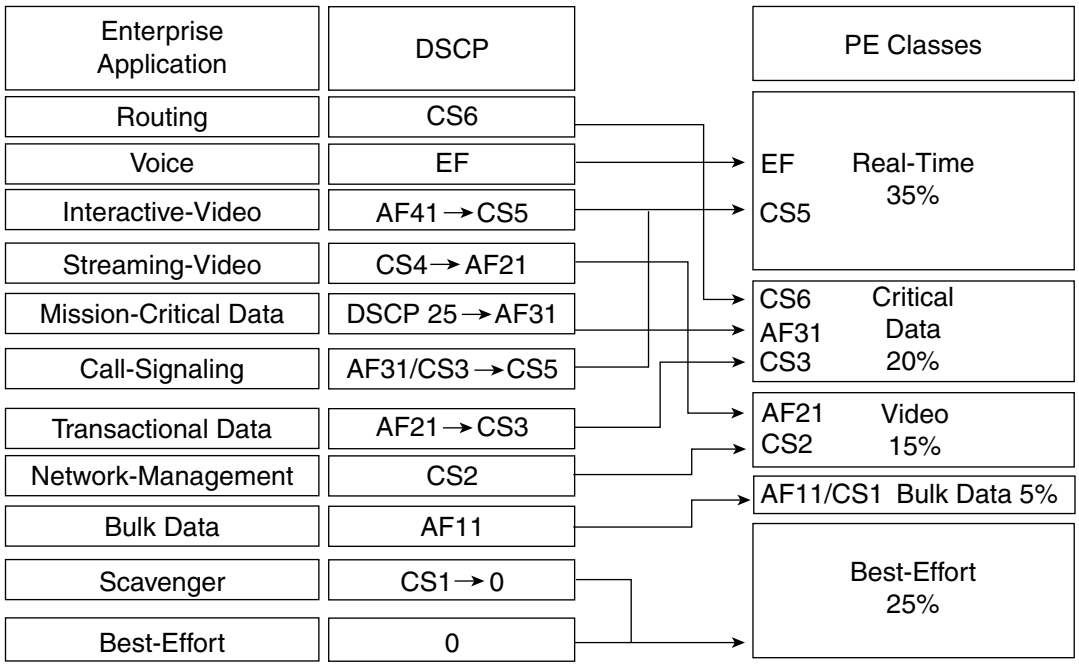
- **show policy**
- **show policy interface**

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

Five-Class Provider-Edge Model: CE Design

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

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



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

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

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

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

Verification commands:

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

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

Provider-Edge QoS Considerations

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

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

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

Service Provider-to-Enterprise Models

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

Three-Class Provider-Edge Model: PE Design

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

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

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

Verification commands:

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

Four-Class Provider-Edge Model: PE Design

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

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

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

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

Verification commands:

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

Five-Class Provider-Edge Model: PE Design

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

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

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

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

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

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

Verification commands:

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

MPLS DiffServ Tunneling Modes

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

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

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

- Uniform Mode
- Short Pipe Mode
- Pipe Mode

Uniform Mode

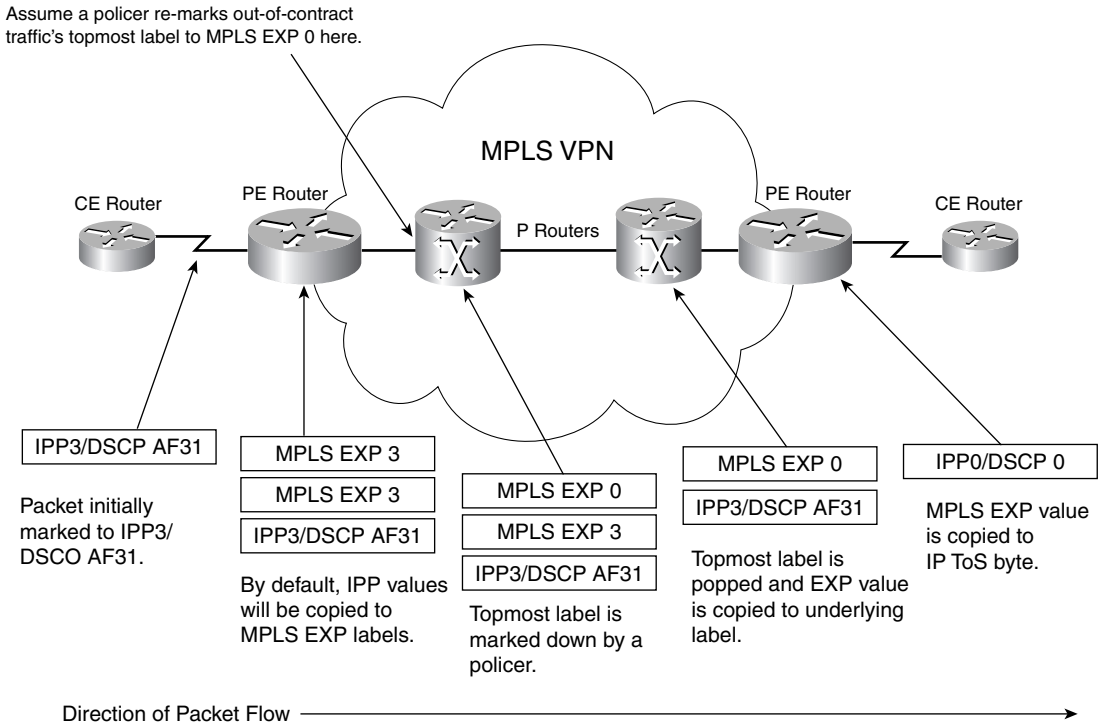
Uniform Mode generally is utilized when the customer and service provider share the same DiffServ domain, as in the case of an enterprise deploying its own MPLS VPN core.

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

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

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

Figure 5-8 MPLS DiffServ Uniform Tunneling Mode Operation

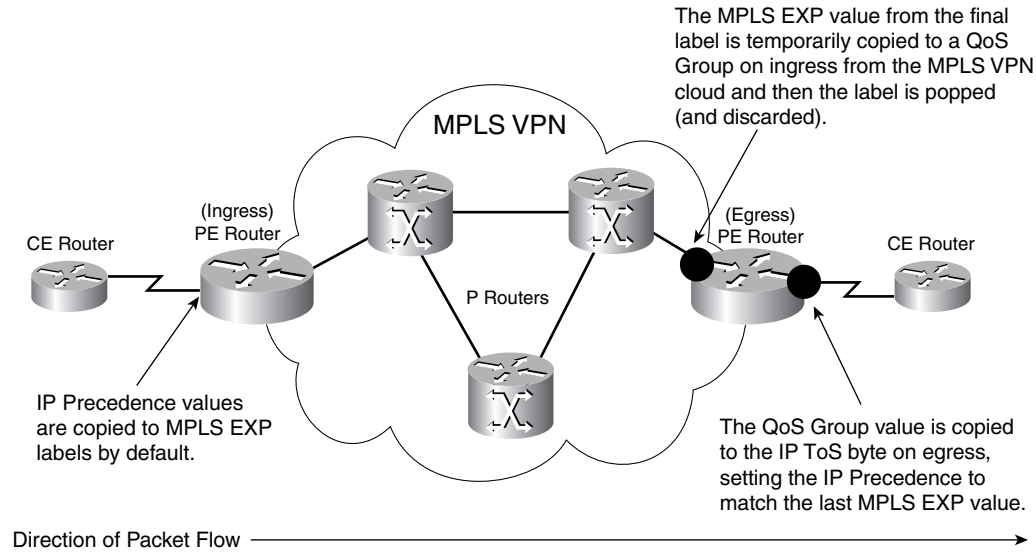


The mapping of IP Precedence to MPLS EXP is performed by default on PEs for customer-to-provider traffic.

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

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

Figure 5-9 MPLS DiffServ Uniform Tunneling Mode Policies



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

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

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

Verification commands:

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

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

Short Pipe Mode

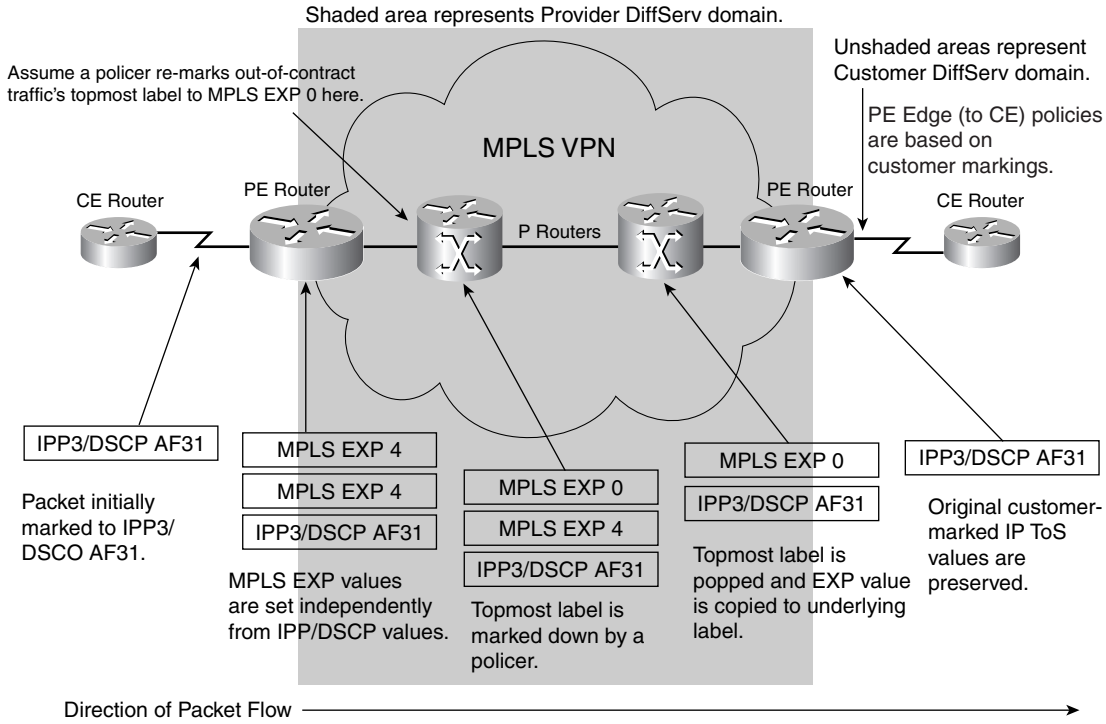
Short Pipe Mode is utilized when the customer and service provider are in different DiffServ domains. (The service provider's DiffServ domain begins at the ingress PE's ingress interface and terminates on the egress PE's ingress interface.)

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

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

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

Figure 5-10 MPLS DiffServ Short Pipe Mode Tunneling Operation



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

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

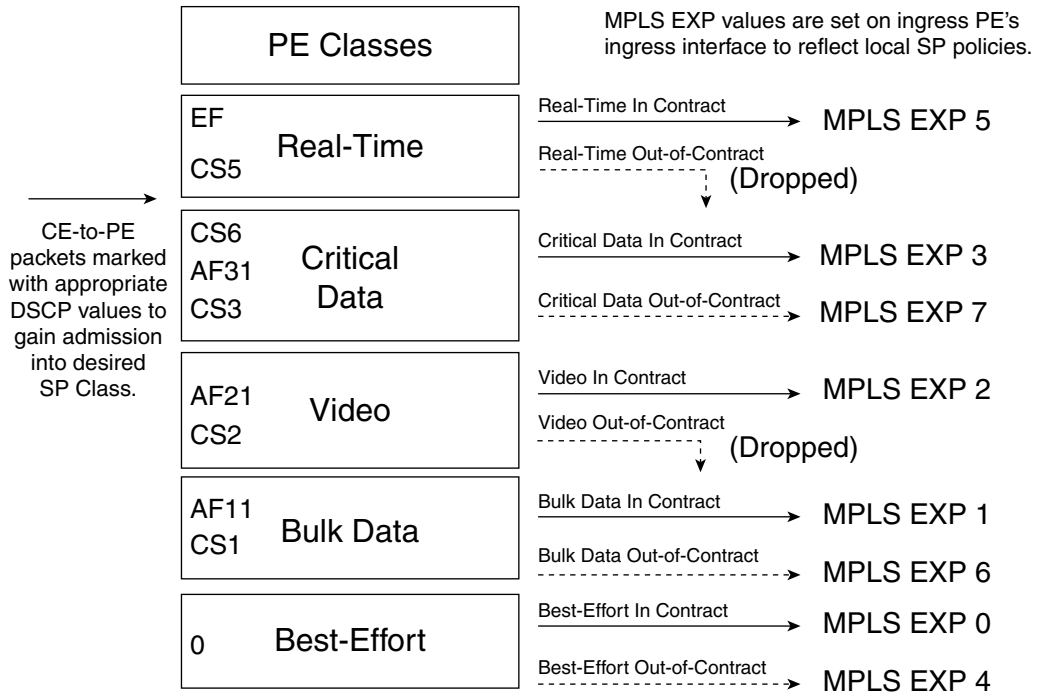
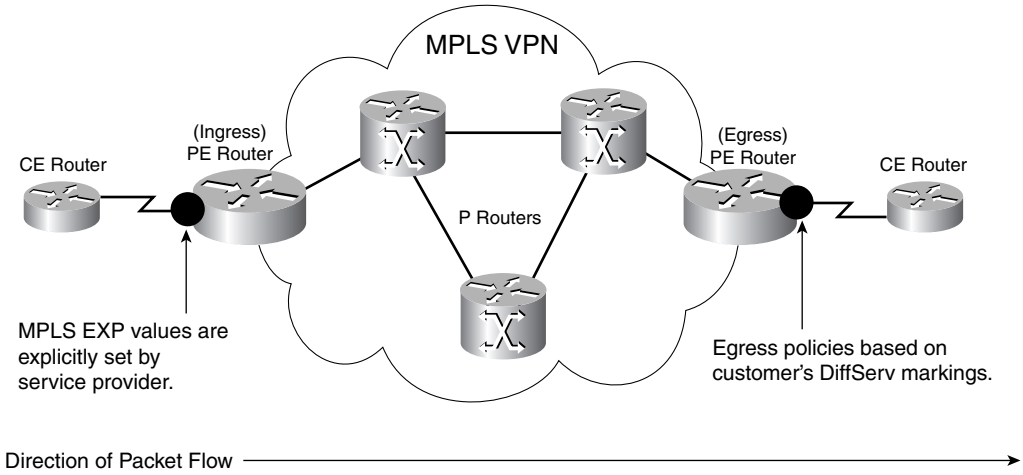


Figure 5-12 shows the ingress PE ingress interface re-marking policies for Short Pipe Mode, based on the re-marking diagram provided in Figure 5-11. No mapping from MPLS EXP to QoS Group is needed on the egress PE's ingress interface (as was required for Uniform Mode) because the MPLS EXP value loses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

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

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

Figure 5-12 MPLS DiffServ Short Pipe Mode Tunneling Policies



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

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

```
!
class-map match-any REALTIME
  match ip dscp ef
  match ip dscp cs5
class-map match-any CRITICAL-DATA
  match ip dscp cs6
  match ip dscp af31
  match ip dscp cs3
class-map match-any VIDEO
  match ip dscp af21
  match ip dscp cs2
class-map match-any BULK-DATA
  match ip dscp af11
  match ip dscp cs1
!
!
 policy-map PE-FIVE-CLASS-SHORT-PIPE-MARKING
  class REALTIME
   police cir 1050000
     conform-action set-mpls-exp-topmost-transmit 5  ! Conforming RT set to 5
     exceed-action drop                              ! Excess Realtime is dropped
  class CRITICAL-DATA
   police cir 600000
     conform-action set-mpls-exp-topmost-transmit 3  ! Critical Data set to 3
     exceed-action set-mpls-exp-topmost-transmit 7   ! Excess Critical set 7
  class VIDEO
   police cir 450000
     conform-action set-mpls-exp-topmost-transmit 2  ! Conforming Video set to 2
     exceed-action drop                              ! Excess Video dropped
  class BULK-DATA
   police cir 150000
```

```
     conform-action set-mpls-exp-topmost-transmit 1  ! Conforming Bulk set to 1
     exceed-action set-mpls-exp-topmost-transmit 6   ! Excess Bulk set to 6
  class class-default
   police cir 750000
     conform-action set-mpls-exp-topmost-transmit 0  ! Conforming BE set to 0
     exceed-action set-mpls-exp-topmost-transmit 4   ! Excess BE set to 4
!
…
!
interface FastEthernet1/0
 description FE TO CUSTOMER RED CE                   ! Link to/from CE
 ip vrf forwarding RED
 ip address 10.1.12.2 255.255.255.0
 service-policy input PE-FIVE-CLASS-SHORT-PIPE-MARKING
!
```

Verification commands:

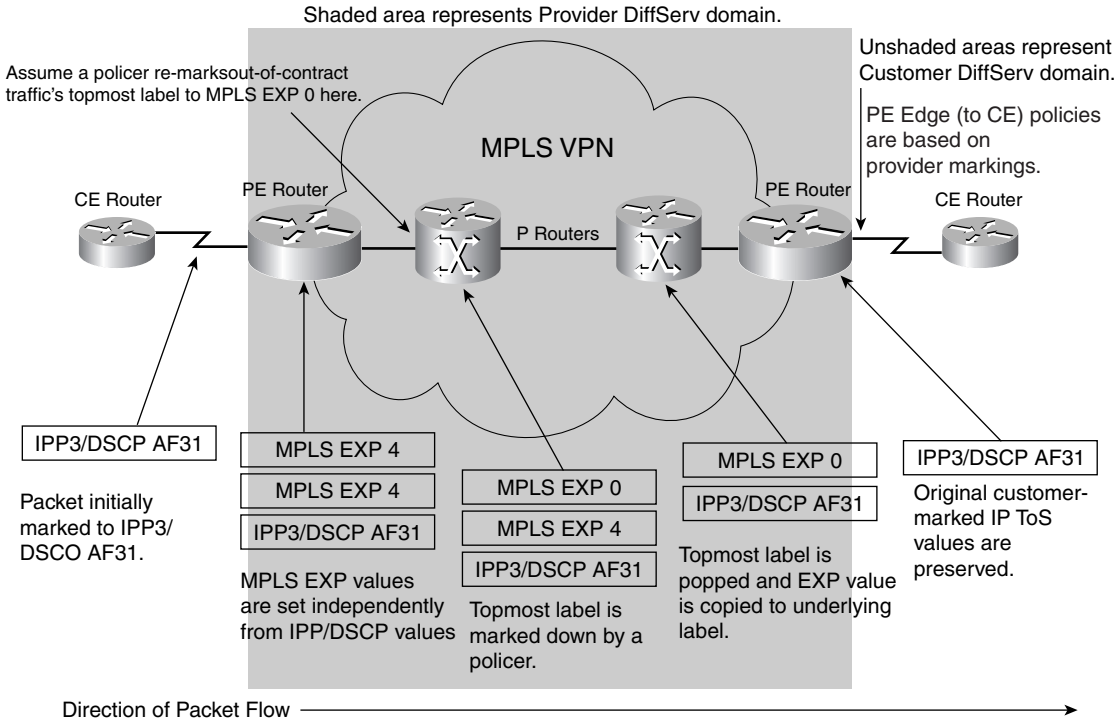
- **show policy**
- **show policy interface**

Pipe Mode

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

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

Figure 5-13 MPLS DiffServ Pipe Mode Tunneling Operation



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

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

Figure 5-14 MPLS DiffServ Pipe Mode Tunneling Policies

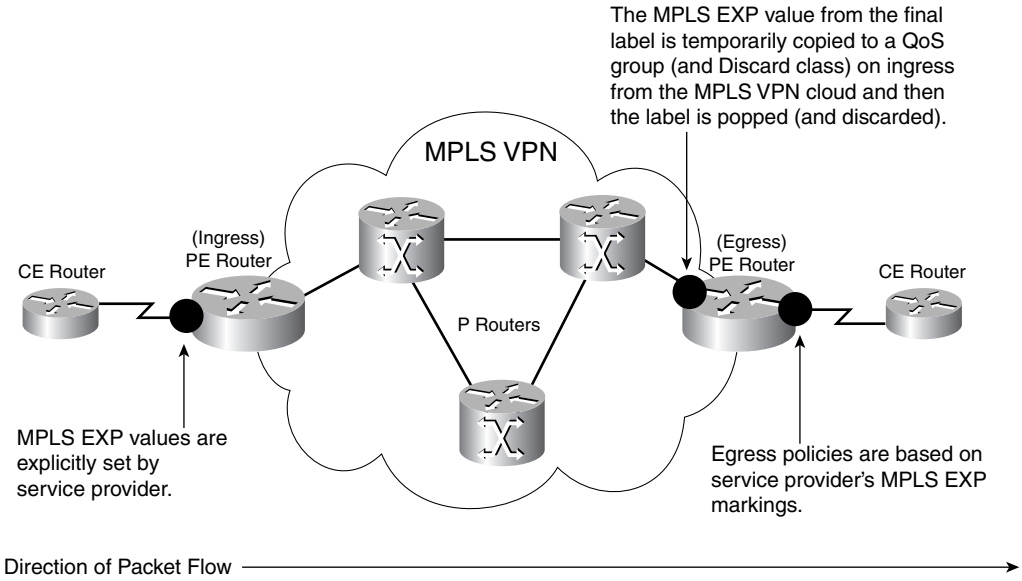
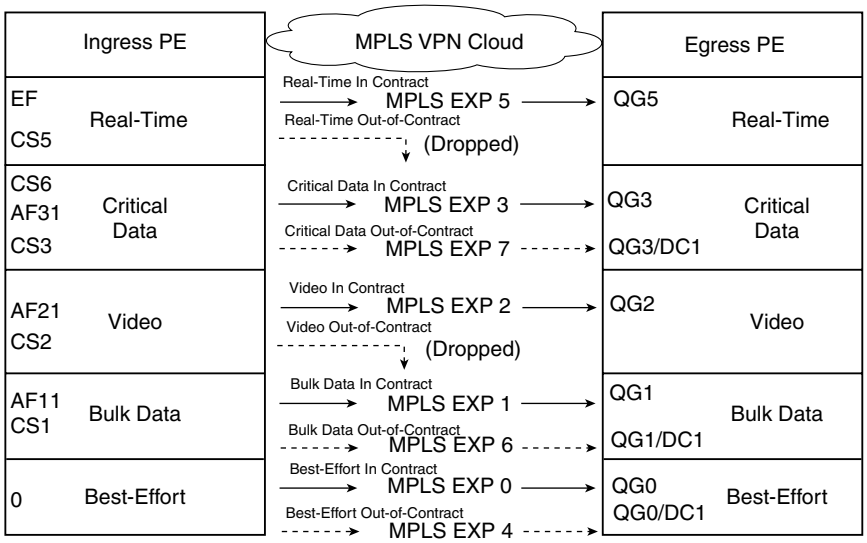


Figure 5-15 illustrates adapting the five-class service provider model to Pipe Mode. The first set of re-markings shows ingress PE re-marking from DSCP to MPLS EXP values, depending on whether the traffic is in contract or out-of-contract. The second set of markings shows how these MPLS EXP values can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

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

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



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

```
!
class-map match-any REALTIME
  match ip dscp ef
  match ip dscp cs5
class-map match-any CRITICAL-DATA
  match ip dscp cs6
  match ip dscp af31
  match ip dscp cs3
class-map match-any VIDEO
  match ip dscp af21
  match ip dscp cs2
class-map match-any BULK-DATA
  match ip dscp af11
  match ip dscp cs1
!
!
 class-map match-all MPLS-EXP-7
  match mpls experimental topmost 7          ! Matches MPLS EXP 7
 class-map match-all MPLS-EXP-6
  match mpls experimental topmost 6          ! Matches MPLS EXP 6
 class-map match-all MPLS-EXP-5
  match mpls experimental topmost 5          ! Matches MPLS EXP 5
 class-map match-all MPLS-EXP-4
  match mpls experimental topmost 4          ! Matches MPLS EXP 4
 class-map match-all MPLS-EXP-3
  match mpls experimental topmost 3          ! Matches MPLS EXP 3
 class-map match-all MPLS-EXP-2
  match mpls experimental topmost 2          ! Matches MPLS EXP 2
 class-map match-all MPLS-EXP-1
  match mpls experimental topmost 1          ! Matches MPLS EXP 1
 class-map match-all MPLS-EXP-0
  match mpls experimental topmost 0          ! Matches MPLS EXP 0
!
!
 class-map match-all QOSGROUP5
  match qos-group 5                          ! Matches QoS Group 5
 class-map match-all QOSGROUP3
```

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

 class MPLS-EXP-7
  set qos-group 3            ! Excess Critical Data is set to QG3
  set discard-class 1        ! Excess Critical Data has DC set to 1
 class MPLS-EXP-2
  set qos-group 2            ! Conforming Video is set to QG 2
 class MPLS-EXP-1
  set qos-group 1            ! Conforming Bulk is set to QG 1
 class MPLS-EXP-6
  set qos-group 1            ! Excess Bulk is set to QG 1
  set discard-class 1        ! Excess Bulk has DC set to 1
 class MPLS-EXP-0
  set qos-group 0            ! Conforming Best Effort is set to QG 0
 class MPLS-EXP-4
  set qos-group 0            ! Excess Best Effort is set to QG 0
  set discard-class 1        ! Excess Best Effort has DC set to 1
!
!
policy-map PE-CE-QUEUING     ! Queuing policy for PE to CE link
 class QOSGROUP5
  priority percent 35        ! Voice class gets 35% LLQ
 class QOSGROUP3
  bandwidth percent 20        ! Critical Data class gets 20% CBWFQ
  random-detect discard-class-based                ! DC-Based WRED is enabled
  random-detect discard-class 0   30    40    10    ! DC 0 is tuned for WRED
  random-detect discard-class 1   20    40    10    ! DC 1 is tuned for WRED
 class QOSGROUP2
```

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

Verification commands:

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

Pipe Mode with an Explicit Null LSP

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

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

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

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

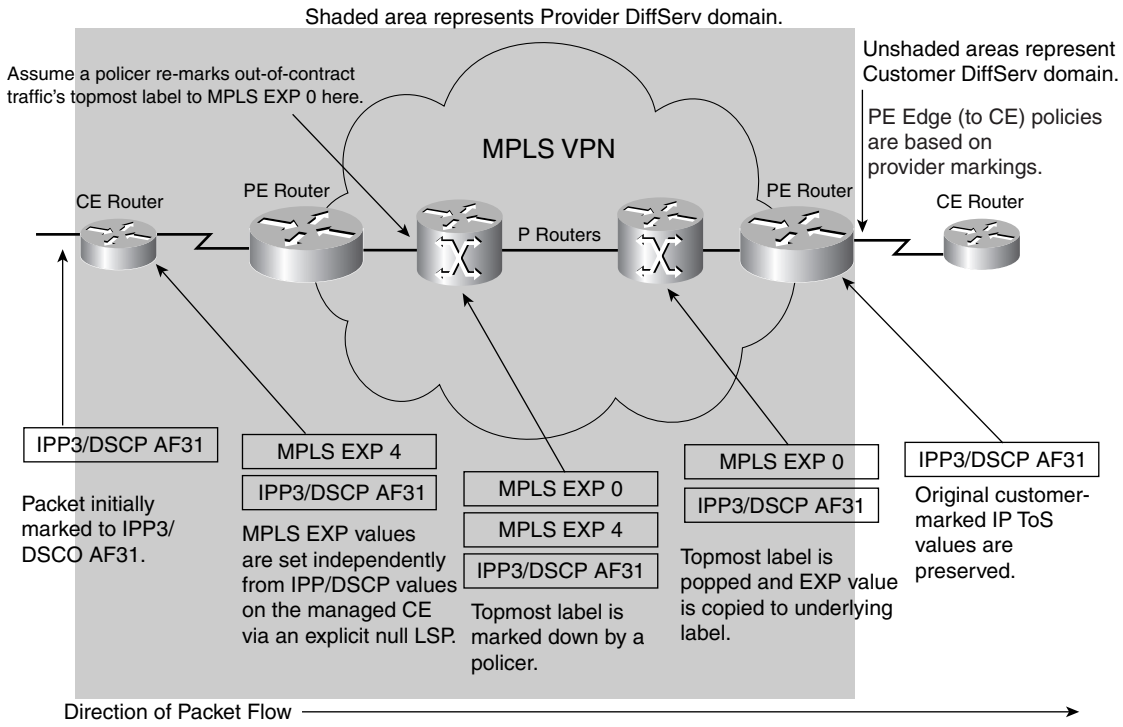
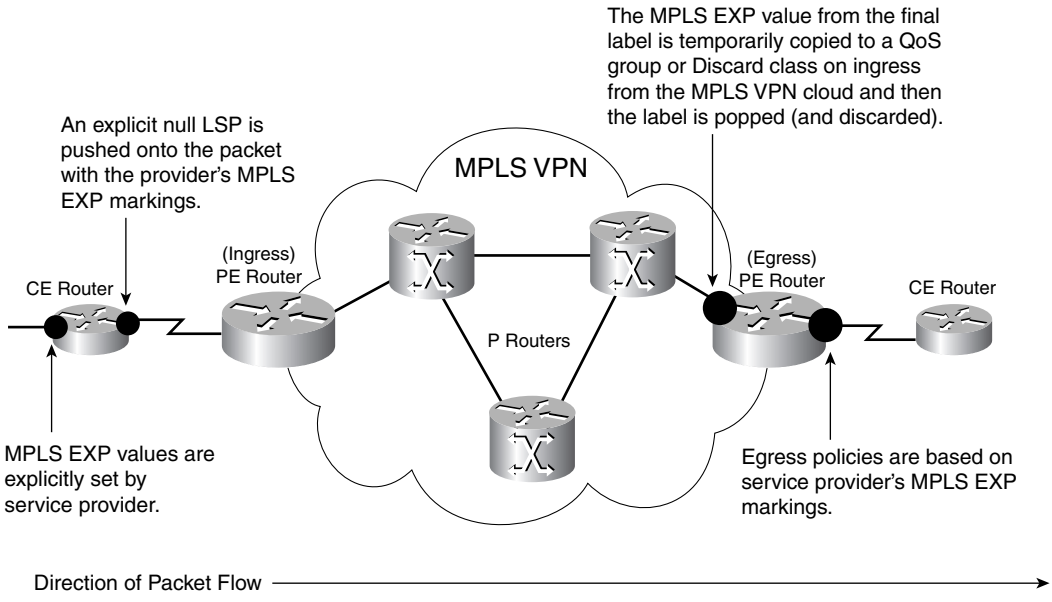


Figure 5-17 shows the points where policies are required for Pipe Mode with Explicit Null LSP MPLS DiffServ tunneling.

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

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



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

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

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

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

Verification commands:

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

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

Summary

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

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

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

References

Standards

- RFC 2547, “BGP/MPLS VPNs”: http://www.ietf.org/rfc/rfc2547.txt
- RFC 2597, “Assured Forwarding PHB Group”: http://www.ietf.org/rfc/rfc2597.txt
- RFC 2702, “Requirements for Traffic Engineering over MPLS”: http://www.ietf.org/rfc/rfc2702.txt

- RFC 2917, "A Core MPLS IP VPN Architecture": http://www.ietf.org/rfc/rfc2917.txt
- RFC 3270, "Multiprotocol Label Switching (MPLS) Support of Differentiated Services": http://www.ietf.org/rfc/rfc3270.txt

Books

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

Cisco Documentation

- Configuring MPLS and MPLS traffic engineering (Cisco IOS Release 12.2): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fswtch_c/swprt3/xcftagc.htm
- MPLS VPNS (Cisco IOS Release 12.2.13T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftvpn13.htm
- MPLS DiffServ tunneling modes (Cisco IOS Release 12.2.13T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftdtmode.htm
- MPLS DiffServ-Aware Traffic Engineering (DS-TE) (Cisco IOS Release 12.2.4T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t4/ft_ds_te.htm
- MPLS Cisco IOS documentation main link (Cisco IOS Release 12.3): http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123cgcr/swit_vcg.htm#999526

CHAPTER 6

IPSec VPN QoS Design

IPSec VPNs are the most widely deployed VPNs and are found in three main contexts:

- Site-to-site IPSec VPNs
- Teleworker IPSec VPNs
- Remote-access client (mobility) IPSec VPNs

QoS considerations for site-to-site and teleworker IPSec VPNs are examined in this design chapter (as QoS is rarely—if ever—deployed in remote-access client IPSec VPN scenarios). These considerations include the following:

- IPSec modes of operation
- Bandwidth and delay increases because of encryption
- IPSec and cRTP incompatibility
- IP ToS byte preservation through IPSec encryption
- QoS and Anti-Replay interaction implications

Following a discussion of these considerations, design recommendations for site-to-site and teleworker (DSL and cable) solutions are presented in detail.

Whereas MPLS technologies provide VPN services, such as network segregation and privacy, by maintaining independent virtual router forwarding tables, IPSec achieves such VPN services through encryption.

As defined in RFCs 2401 through 2412, IPSec protocols provide mechanisms to enable remote sites to cost effectively connect to enterprise intranets or extranets using the Internet (or a service provider's shared IP networks). Because of IPSec protocol encryption, such VPNs can provide the same management and security policies as private networks.

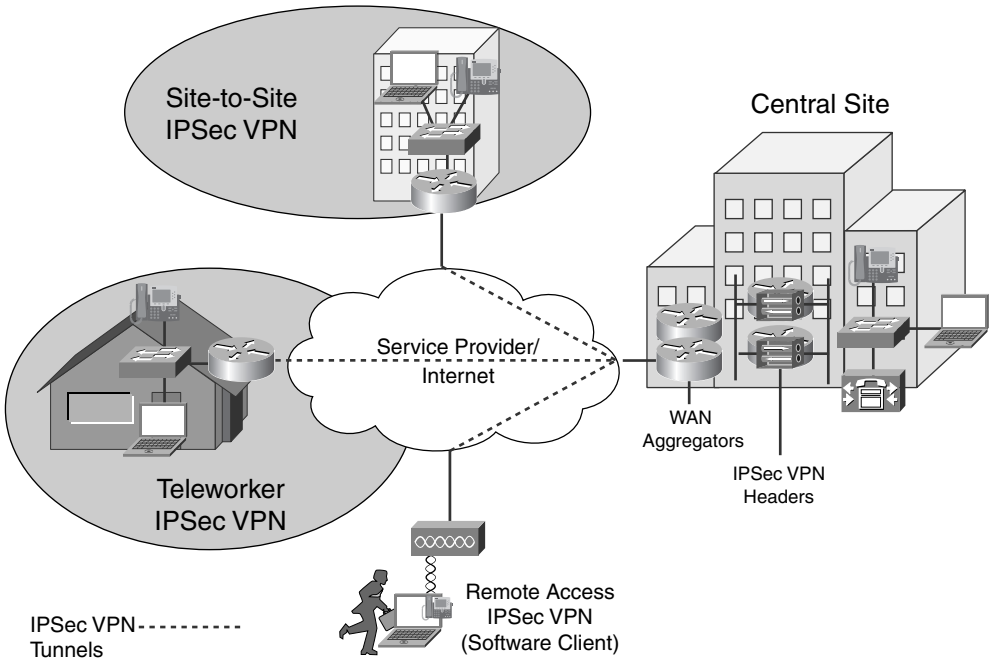
IPSec VPN services are built by overlaying a point-to-point mesh over the Internet using Layer 3–encrypted tunnels. This architecture requires security appliances, such as (hardware or software) firewalls or routers that support IPSec tunnels, to be installed at both ends of each tunnel. Encryption/decryption is performed at these tunnel endpoints, and the protected traffic is carried across the shared network.

Three main design contexts for IPSec VPNs exist, as shown in Figure 6-1:

- **Site-to-site VPNs**—Tunnels are maintained by hardware devices to connect multiple users at remote branches to (one or more) central sites.
- **Teleworker VPNs**—Tunnels are maintained by hardware devices to connect (typically) a single user at his or her residence to a central site.

- **Remote-access clients**—Tunnels are established by software to connect mobile users at airports, hotels, or similar places to a central site using WLAN hotspots, LAN ports, or modems.

Figure 6-1 IPSec VPN Design Contexts



Enabling converged services, such as voice and video, on an IPSec VPN has been dubbed V3PN. V3PN is essentially the overlaying of QoS technologies over IPSec VPNs to provide the required service levels to voice and video applications. As such, V3PN solutions relate to only two of the three IPSec VPN design contexts: site-to-site VPNs and telecommuter VPNs. (Little, if any, QoS is available in remote-access client networks.)

This chapter discusses QoS design considerations and recommendations for both site-to-site and telecommuter V3PN solutions.

Note It is beyond the scope of this chapter to detail IPSec encryption operation and configuration; a working knowledge of IPSec is assumed.

Site-to-Site V3PN QoS Considerations

Attractive pricing is usually the driver behind deploying site-to-site IPSec VPNs as an alternative to private WAN technologies. Many of the same considerations required by private WANs need to be taken into account for IPSec VPN scenarios because they usually are deployed over the same Layer 2 WAN access media.

IPSec VPNs also share some similar concerns with MPLS VPNs. For instance, the enterprise's end-to-end delay and jitter budgets depend significantly on the service provider's SLAs. Therefore, enterprises deploying V3PN solutions are recommended to utilize Cisco Powered Network IP Multiservice service providers, as discussed in Chapter 5, "MPLS VPN QoS Design."

However, IPSec VPNs present many unique considerations for QoS design, including the following (each is discussed in detail throughout the rest of the chapter):

- IPSec VPN modes of operation
- Packet overhead increases because of encryption
- cRTP and IPSec incompatibility
- Prefragmentation
- Bandwidth provisioning
- Logical topologies
- Delay budget increases because of encryption
- ToS byte preservation
- QoS Pre-Classify feature
- Pre-encryption queuing
- Anti-Replay implications
- Control plane provisioning

IPSec VPN Modes of Operation

Three principal modes of IPSec VPN operation exist:

- IPSec Tunnel Mode (No IP GRE Tunnel)
- IPSec Transport Mode with an Encrypted IP GRE Tunnel
- IPSec Tunnel Mode with an Encrypted IP GRE Tunnel

The advantages, disadvantages, features, and limitations of these options are discussed next.

IPSec Tunnel Mode (No IP GRE Tunnel)

This option does not utilize an IP GRE tunnel. With this option, only IPSec unicast traffic can be transported. (IP multicast traffic cannot be transported between IPSec peers without configuring an IP GRE tunnel.)

This configuration might be sufficient to support application requirements; its advantage lies in lower CPU overhead (primarily at the headend IPSec VPN router) compared with alternative IPSec design options.

IPSec security associations (SAs) are created for each access list line matched. An access list must be specified in the crypto map to designate packets that are to be encrypted. Such an access list typically entails several lines to define the application(s) to be encrypted by the five ACL tuples: source/destination IP address, protocol, and source/destination port numbers. When not encrypting a GRE tunnel, it is possible to create a separate SA for each application or access-list line match or to create an SA that carries all traffic that matches an ACL range (which is recommended). Each SA has its own Encryption Security Protocol (ESP) or Authentication Header (AH) sequence number.

Anti-Replay drops can be eliminated or minimized by constructing access lists that create a separate security association for each class of traffic being influenced by per-hop QoS policies. (Anti-Replay is an IPSec standard feature that discards packets that fall outside a receiver's 64-byte sliding window because such packets are considered suspect or potentially compromised—it is discussed in greater detail later in this chapter.)

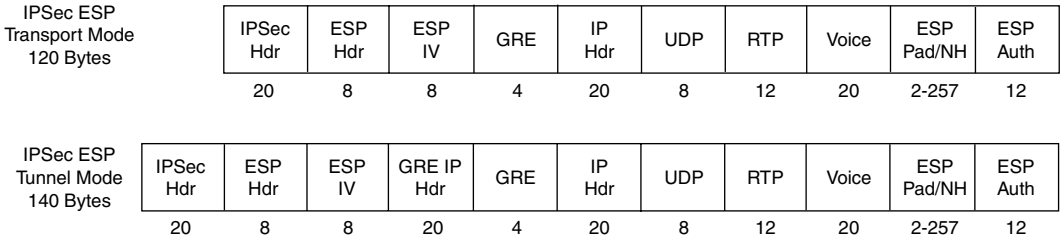
The Cisco IOS feature of prefragmentation for IPSec VPNs (also discussed later in this chapter) is supported in IPSec tunnel mode (no IP GRE tunnel) as of Cisco IOS Release 12.2(12)T.

IPSec Transport Mode with an Encrypted IP GRE Tunnel

IPSec transport mode (encrypting an IP GRE tunnel) is a commonly deployed option because it provides all the advantages of using IP GRE, such as IP Multicast protocol support (and, thus, also the support of routing protocols that utilize IP Multicast) and multiprotocol support. Furthermore, this option saves 20 bytes per packet over IPSec tunnel mode (encrypting an IP GRE tunnel) because an additional IP header is not required. Figure 6-2 illustrates IPSec transport mode versus tunnel mode when encryption is performed in an IP GRE tunnel.

The IPSec peer IP addresses and the IP GRE peer address must match for transport mode to be negotiated; if they do not match, tunnel mode is negotiated.

Figure 6-2 IPSec Transport Mode Versus Tunnel Mode for a G.729 VoIP Packet



The Cisco IOS prefragmentation feature for IPSec VPNs (discussed later in the chapter) is *not* supported for transport mode because the decrypting router cannot determine whether the fragmentation was done before or after encryption (for example, by a downstream router between the encrypting and decrypting routers).

Although IPSec transport mode saves a small to moderate amount of link bandwidth, it does not provide any reduction in packets per second switched by the router. Therefore, because the number of packets per second primarily affects CPU performance, no significant CPU performance gain is realized by using IPSec transport mode.

IPSec tunnel mode is the default configuration option. To configure transport mode, it must be specified under the IPSec transform set, as shown in Example 6-1.

Example 6-1 Enabling IPSec Transport Mode

```
!
crypto ipsec transform-set ENTERPRISE esp-3des esp-sha-hmac
 mode transport             ! Enables IPSec Transport mode
!
```

IPSec Tunnel Mode with an Encrypted IP GRE Tunnel

IPSec tunnel mode (encrypting an IP GRE tunnel) is the primarily recommended IPSec VPN design option. Although it incurs the greatest header overhead of the three options, it is capable of supporting IP Multicast (with the capability to run a dynamic routing protocol within the IP GRE tunnel for failover to an alternative path), and it supports prefragmentation for IPSec VPNs.

When configured with a routing protocol running within an IP GRE tunnel, the routing protocol's Hello packets maintain the security associations between the branch and both (assuming a redundant configuration) headend routers. There is no need to create a security association with a backup headend peer if the primary peer fails.

Note The design principles in this chapter were proven by scalability testing in the Cisco Enterprise Solutions Engineering labs. These large-scale testing methods were designed to test worst-case scenarios. From a design standpoint, these entailed enabling the following:
— Strong Triple-Digital Encryption Standard (3DES) encryption for both Internet Key Exchange (IKE) and IPSec
— IP GRE with IPSec tunnel mode
— Diffie-Hellman Group 2 (1024 bit) for IKE
— Secure Hash Algorithm (SHA) 160-bit RFC 2104 Keyed-Hashing for Message Authentication (HMAC) with RFC 1321 Message Digest 5 (MD5)
— (MD5)-HMAC (both hash algorithms truncated to 12 bytes in the ESP packet trailer)
— Preshared keys
If an enterprise chooses to implement less stringent security parameters, to use IPSec transport mode instead of tunnel mode, or to not implement IP GRE tunnels, the designs continue to be applicable from functional and scalability standpoints.

Packet Overhead Increases

The addition of tunnel headers and encryption overhead increases the packet sizes of all encrypted applications: voice, video, and data. This needs to be taken into account when provisioning LLQ or CBWFQ bandwidth to a given class.

For example, consider voice. The two most widely deployed codecs for voice are G.711 and G.729. Each codec typically is deployed at 50 pps (generating 20-ms packetization intervals).

The Layer 3 data rate for a G.711 call (at 50 pps) is 80 kbps. IP Generic Routing Encapsulation (GRE) tunnel overhead adds 24 bytes per packet. The IPSec Encapsulating Security Payload (ESP) adds another 56 bytes. The combined additional overhead increases the rate from 80 kbps (clear voice) to 112 kbps (IPSec ESP tunnel-mode encrypted voice).

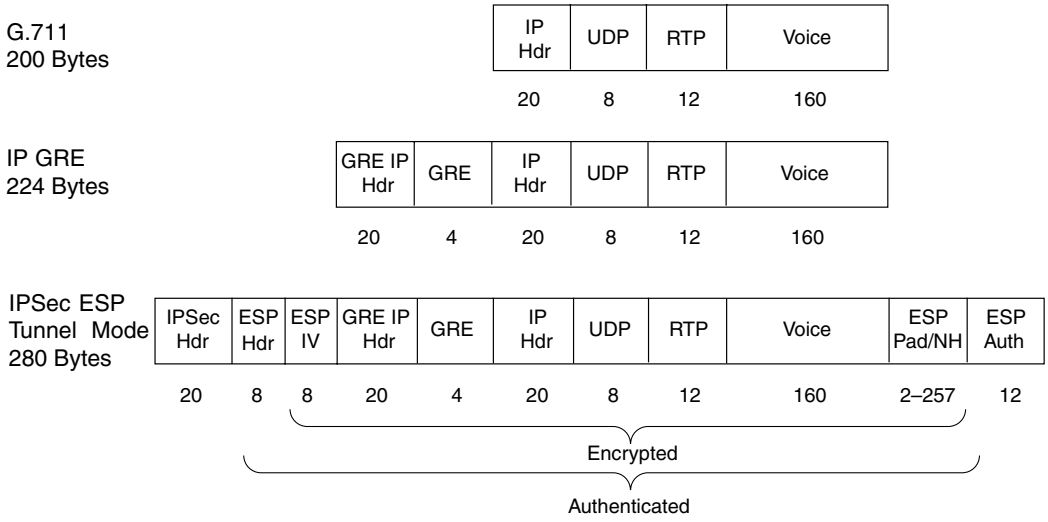
The calculation is as follows:

```
    200 bytes per packet (G.711 voice)
     24 bytes per packet (IP GRE overhead)
+    56 bytes per packet (IPSec ESP overhead)
    280 bytes per packet
·     8 bits per byte
   2240 bits per packet
·    50 packets per second
 112,000 bits per second
```

The additional overhead represents a 40 percent increase in the bandwidth required for an encrypted G.711 call.

The 280-byte packet's header, data, and trailer fields for an IPSec tunnel-mode ESP encrypted G.711 call are shown in Figure 6-3.

Figure 6-3 Anatomy of an IPSec-Encrypted G.711 Packet



The Layer 3 data rate for a G.729 call (at 50 pps) is 24 kbps. IP GRE tunnel overhead adds 24 bytes per packet. IPSec ESP adds another 52 bytes. The combined additional overhead increases the rate from 24 kbps (clear voice) to just less than 56 kbps (IPSec ESP tunnel-mode encrypted voice).

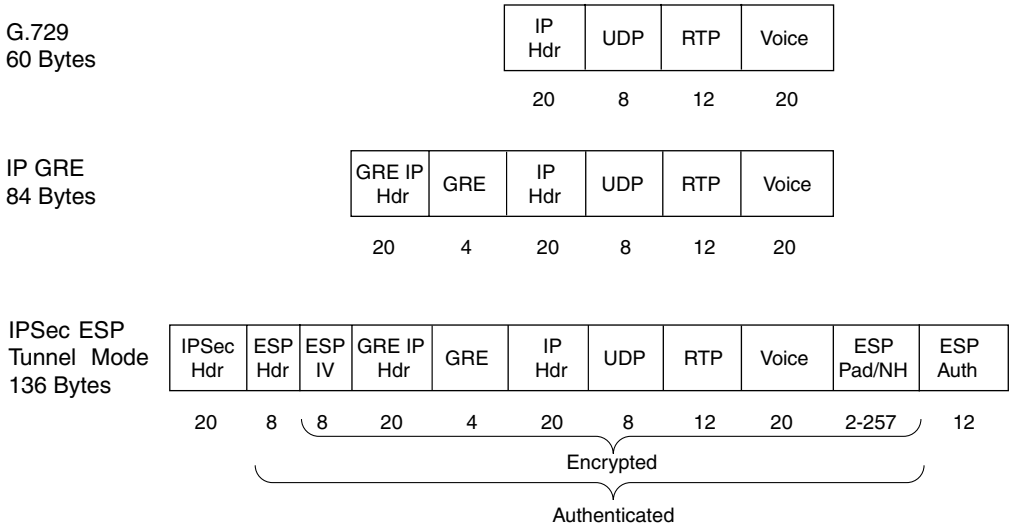
The calculation is as follows:

60 bytes per packet (G.729 voice)
24 bytes per packet (IP GRE overhead)
\+ 52 bytes per packet (IPSec ESP overhead)
136 bytes per packet
· 8 bits per byte
1088 bits per packet
· 50 packets per second
54,400 bits per second

The additional overhead represents a 227 percent increase in the bandwidth required for an encrypted G.729 call.

The 136-byte packet's header, data, and trailer fields for an IPSec tunnel-mode ESP encrypted G.729 call are shown in Figure 6-4.

Figure 6-4 Anatomy of an IPSec-Encrypted G.729 Packet



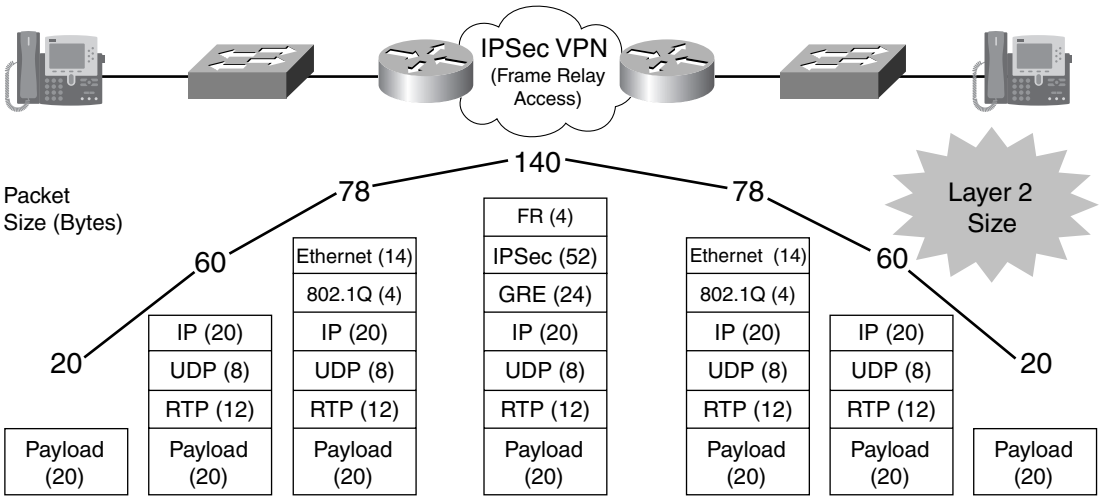
It is important to note that these bandwidth allocations are Layer 3 bandwidth requirements and *do not include* Layer 2 overhead (which is media dependent). Therefore, Layer 2 overhead needs to be added on top of the Layer 3 requirements in provisioning LLQ and CBWFQ bandwidth. This is illustrated in Figure 6-5, where Ethernet overhead (Ethernet plus 802.1Q trunking) and Frame Relay overhead are added and removed from the packet in transit.

Key Layer 2 overhead values are reiterated in Table 6-1.

Table 6-1 Layer 2 Encapsulation Overhead

Layer 2 Encapsulation	Overhead
Ethernet	14 bytes (+ 4 for 802.1Q)
Frame Relay	4 bytes (+ 4 for FRF.12)
MLP	10 bytes (+ 3 for MLP LFI)
ATM	5 bytes per 53-byte cell + cell padding (variable)

Figure 6-5 Packet Size Changes of a G.729 IPSec-Encrypted Packet



Therefore, the calculation, inclusive of Layer 2 overhead, is as follows. This example assumes that a G.729 call will be encrypted over a slow speed (£ 768-kbps Frame Relay link), which requires FRF.12 fragmentation and interleaving.

60 bytes per packet (G.729 voice)
24 bytes per packet (IP GRE overhead)
52 bytes per packet (IPSec ESP overhead)
4 bytes per packet (FR overhead)
\+ **4 bytes per packet (FRF.12 overhead)**
44 bytes per packet
· 8 bits per byte
1152 bits per packet
· 50 packets per second
57,600 bits per second (rounded up to 58 kbps)

In summary, it is important always to include Layer 2 overhead in accurate bandwidth provisioning for IPSec-encrypted applications.

cRTP and IPSec Incompatibility

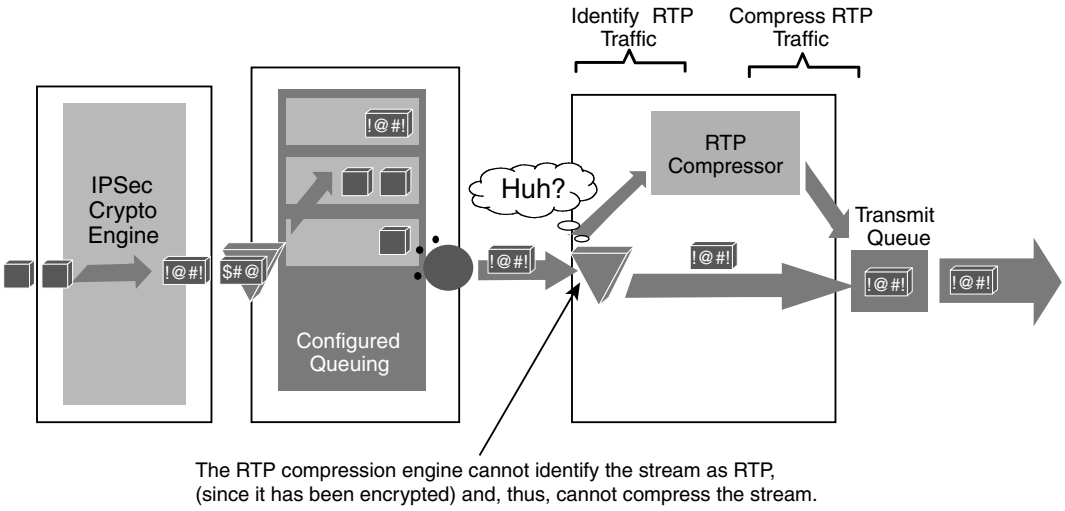
The significant increases in bandwidth required by IPSec encryption lead many administrators to consider the use of IP RTP header compression (cRTP) to offset these increases.

However, one of the caveats of encryption is that key portions of the original IP packet that could be referenced for QoS (and other) purposes are no longer readable. Such is the case with cRTP.

cRTP and IPSec are inherently incompatible standards. The original IP/UDP/RTP header already is encrypted by IPSec by the time the RTP compressor is called upon to perform the compression. Therefore, because cRTP cannot associate the encrypted IP/UDP/RTP packet with a known media stream, compression cannot occur and cRTP bandwidth savings cannot be realized. The encrypted IP/UDP/RTP packet simply bypasses the compression process and continues (uncompressed) to the transmit queue.

This is illustrated in Figure 6-6.

Figure 6-6 IPSec and cRTP Incompatibility



It is important to recognize that cRTP functions on a hop-by-hop basis, whereas IPSec can span multiple intermediate (Layer 3) hops between IPSec endpoints. This distinction further exacerbates incompatibility between the features.

Although developments are under way to address these incompatibilities, at the time of this writing, cRTP cannot be utilized to achieve bandwidth savings in an IPSec VPN environment.

Prefragmentation

A problem arises when a packet is nearly the size of the maximum transmission unit (MTU) of the outbound link of the encrypting router and then is encapsulated with IPSec headers. The resulting packet is likely to exceed the MTU of the outbound link. This causes packet fragmentation after encryption, which makes the decrypting router reassemble in the process path.

Cisco IOS Release 12.2(13)T introduced a new feature: prefragmentation for IPSec VPNs. Prefragmentation increases the decrypting router's performance by enabling it to operate in the high-performance CEF path instead of the process path.

This feature enables an encrypting router to predetermine the encapsulated packet size from information available in transform sets, which are configured as part of the IPSec security association. If it is predetermined that the packet will exceed the MTU of the output interface, the packet is fragmented before encryption. This function avoids process-level reassembly before decryption and helps improve decryption performance and overall IPSec traffic throughput.

Prefragmentation for IPSec VPNs is enabled globally by default for Cisco VPN routers running Cisco IOS Release 12.2(13)T or higher.

Bandwidth Provisioning

Chapter 1, "Quality of Service Design Overview," presented the 33 Percent LLQ Rule, along with the design rationale behind the recommendation. Furthermore, the rule was expressed as a conservative design recommendation that might not be valid under all constraints. Provisioning for VoIP over IPSec on slow links sometimes poses constraints that might preclude applying the 33 Percent LLQ Rule.

As shown in Table 6-2, the percentage of LLQ required on 64-, 128-, and 256-kbps links for a single encrypted G.729 call exceeds the recommended 33 percent LLQ limit. Enterprises considering such deployments must recognize the impact on data traffic traversing such links when encrypted voice calls were made—specifically, data applications would slow down significantly. If that is considered an acceptable trade-off, not much can be said or done. Otherwise, it is recommended to increase bandwidth to the point that encrypted VoIP calls can be made and still fall within the 33 percent bandwidth recommendation for priority queuing.

When planning the bandwidth required for a branch office, consider the number of concurrent calls traversing the IPSec VPN that the branch is expected to make during peak call periods. This varies based on the job function of the employees located at a branch. For example, an office of software engineers would be expected to make fewer calls than an office of telemarketers. A typical active call ratio may be one active call for every six people (1:6), but this could range from 1:4 or 1:10, depending on the job function of the employees. Given the 512-kbps link from Table 6-2 as an example, with a target of 3 G.729 calls, this link theoretically could support a branch office of between 12 and 30 people. As with all other topologies, call admission control must be administered properly to correspond to the QoS policies deployed within the network infrastructure.

Table 6-2 G.729 Calls by Link Speeds (FRF.12 Is Enabled on Link Speeds ≤ 768 kbps Only)

Line Rate (kbps)	Maximum Number of G.729 Calls	LLQ Bandwidth (kbps)	LLQ Bandwidth (Percentage)
64 (FRF.12)	1 (58 kbps)	58	91%
128 (FRF.12)	1 (58 kbps)	58	46%
256 (FRF.12)	2 (58 kbps)	116	46%
512 (FRF.12)	3 (58 kbps)	174	34%
768 (FRF.12)	4 (58 kbps)	232	31%
1024	6 (56 kbps)	336	33%
1536	9 (56 kbps)	504	33%
2048	12 (56 kbps)	672	33%

Note Although VoIP has been discussed as a primary example of bandwidth provisioning considerations when deploying QoS over IPSec VPNs, it is important to recognize that VoIP is not the only application that might require such considerations; this exercise needs to be performed for *any* application that is being encrypted.

Unlike VoIP, however, other applications—such as video and data—have varying packet rates and sizes. Therefore, crisp provisioning formulas might not apply. Moderate traffic analysis and best guesses, along with trial-and-error tuning, are usually the only options for factoring bandwidth provisioning increases for non-VoIP applications.

Logical Topologies

Similar to private WANs, but unlike fully meshed MPLS VPNs, IPSec VPNs typically are deployed in a (logical) hub-and-spoke topology.

The QoS implications of hub-and-spoke topologies include that access rates to remote sites need to be constrained and shaped at the hub, to avoid delays and drops within the provider's cloud. Shaping at the IPSec VPN hub is done in a manner similar to that of private WAN NBMA media, such as Frame Relay or ATM. Refer to the Headend VPN Edge QoS Options for Site-to-Site V3PNs section, later in this chapter, and also to Chapter 3, "WAN Aggregator QoS Design."

IPSec VPNs are not limited to hub-and-spoke designs. They also can be deployed in partial-mesh or even fully meshed topologies. In such cases, shaping is recommended on any links where speed mismatches occur (similar to private WAN scenarios).

Another alternative is to deploy IPSec VPNs via Dynamic Multipoint Virtual Private Networks (DMVPN), which establish site-to-site IPSec tunnels as needed and tear them down when they no longer are required. As with the previously discussed logical topologies, shaping is required on DMVPN NBMA links with speed mismatches. Specifically, shapers are required to be created dynamically and applied to *logical* DMVPN tunnels to offset any speed mismatches attributed to *physical* NMBA links. However, as of the time of this writing, no shaping or queuing solution exists to guarantee QoS SLAs over DMVPN topologies (although Cisco IOS solutions currently are being evaluated and are in development).

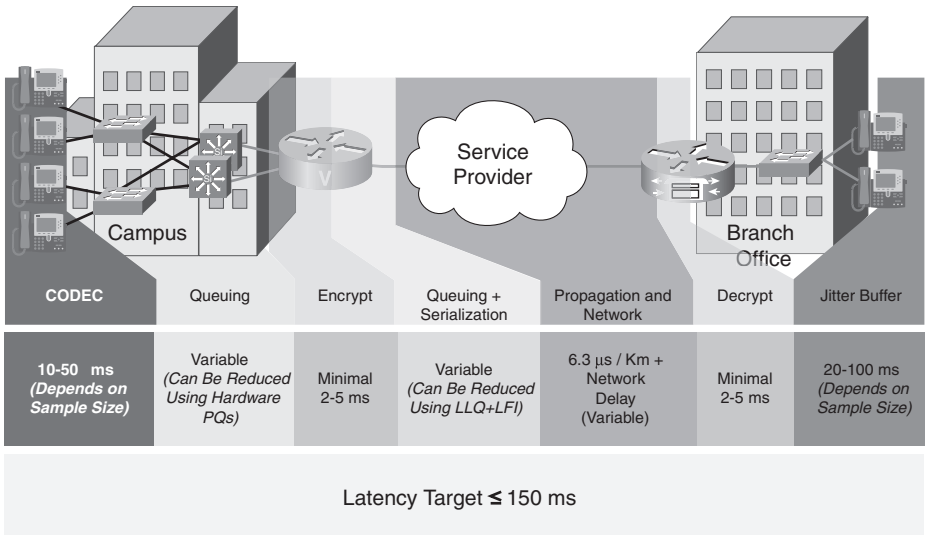
Delay Budget Increases

As previously discussed, the delay budget for a typical IP Telephony implementation includes fixed and variable components. The ITU G.114 specification's target value for one-way delay is 150 ms. In an IPSec VPN deployment, however, two additional delay components must be factored into the overall delay budget:

- Encryption delay at the origination point of the IPSec VPN tunnel
- Decryption delay at the termination point of the IPSec VPN tunnel

Performance and scalability testing results suggest that, in most cases, the additional delay caused by encryption and decryption is approximately 4 to 10 ms (combined). These incremental delays are shown in Figure 6-7.

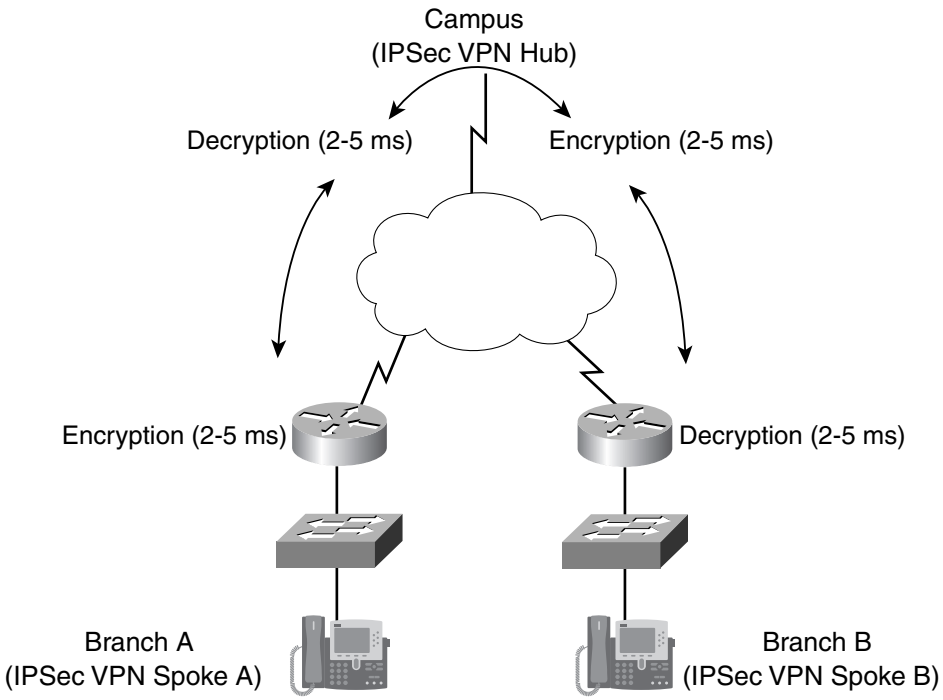
Figure 6-7 IPSec Encryption/Decryption Incremental Delays



A conservative planning estimate would be 10 ms for encryption delay and 10 ms for decryption delay.

This delay might not seem significant for a campus-to-branch call (hub-to-spoke), but the delay might be more relevant in branch-to-branch (spoke-to-spoke) scenarios because encryption and decryption might occur twice (depending on the logical topology of the VPN). This is illustrated in Figure 6-8.

Figure 6-8 IPSec VPN Spoke-to-Spoke Encryption/Decryption Delays



Note Not only do encryption delays need to be factored into spoke-to-spoke IPSec VPN scenarios, but queuing and serialization delays for both legs of the tunnel do as well (as they also would in private WAN spoke-to-spoke scenarios).

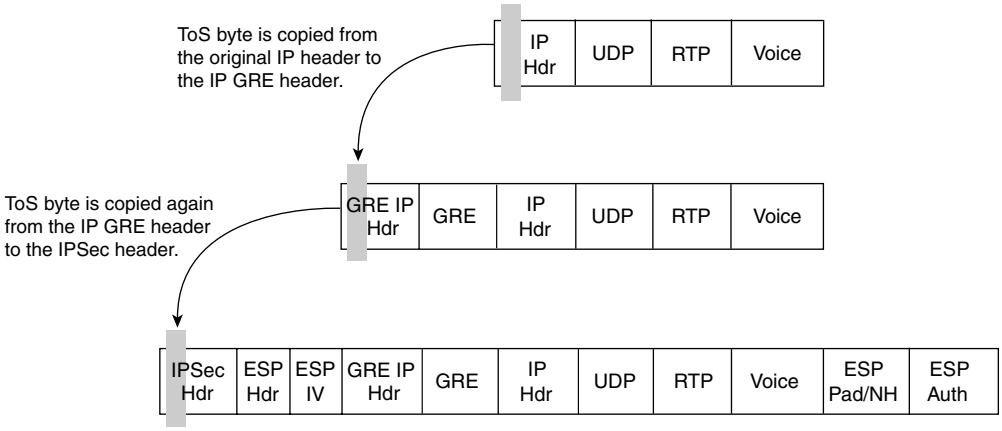
ToS Byte Preservation

For the majority of QoS designs discussed thus far, classification is performed based on DSCP markings in the ToS byte of the IP packet. However, when an IP packet is encrypted through IPSec, the original ToS byte values also are encrypted and, thus, unusable by QoS mechanisms that process the packet (post encryption).

To overcome this predicament, the IPSec protocol standards inherently have provisioned the capability to preserve the ToS byte information of the original IP header by copying it to the IP headers added by the tunneling and encryption process.

As shown in Figure 6-9, the original IP ToS byte values are copied initially to the IP header added by the GRE encapsulation. Then these values are copied again to the IP header added by IPSec encryption.

Figure 6-9 IP ToS Byte Preservation



This process compensates for the fact that the original IP header (including the ToS byte) is actually unreadable (because of encryption) and allows the packet to be processed by (post encryption) QoS mechanisms in the same manner as any other packet.

Additionally, this process underscores the importance of ensuring that the encrypted traffic is marked properly (at Layer 3) before encryption.

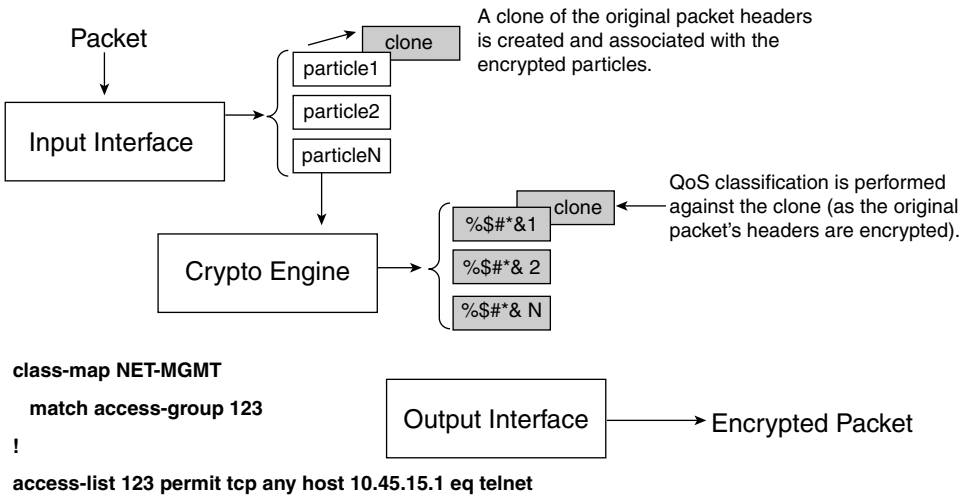
QoS Pre-Classify

The QoS Pre-Classify feature often is confused with ToS byte preservation. QoS Pre-Classify is a Cisco IOS feature that allows for packets to be classified on header parameters other than ToS byte values after encryption.

Because all original packet header fields are encrypted, including source or destination IP addresses, Layer 4 protocol, and source or destination port addresses, post-encryption QoS mechanisms cannot perform classification against criteria specified within any of these fields.

A solution to this constraint is to create a clone of the original packet's headers before encryption. The crypto engine encrypts the original packet, and then the clone is associated with the newly encrypted packet and sent to the output interface. At the output interface, any QoS decisions based on header criteria, except for ToS byte values—which have been preserved—can be performed by matching on any or all of the five access-list tuple values of the clone. In this manner, advanced classification can be administered even on encrypted packets. The process is illustrated in Figure 6-10.

Figure 6-10 QoS Pre-Classify Feature Operation



A key point to remember regarding QoS Pre-Classify is that it is applicable only at the encrypting router's output interface. The fields preserved by QoS Pre-Classify are not available to any routers downstream; the clone never leaves the router performing the encryption, thus ensuring the integrity and security of the IPSec VPN tunnel.

QoS Pre-Classify is supported in all Cisco IOS switching paths and is recommended to be enabled on some platforms even when only the ToS byte values are being used for classification. Testing has shown that when hardware-based encryption cards are combined with QoS, the Cisco IOS Software implementation of the QoS Pre-Classify feature slightly enhances performance, even when matching only on ToS byte values. Furthermore, enabling QoS Pre-Classify by default eliminates the possibility that its configuration will be overlooked if the QoS policy later is changed to include matching on IP addresses, ports, or protocols.

Design recommendations for the QoS Pre-Classify feature can be summarized as follows:

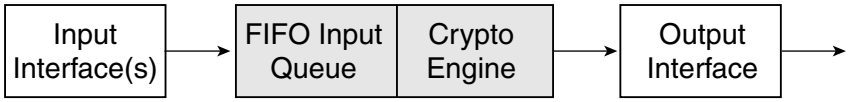
- Enable QoS Pre-Classify on all branch IPSec VPN routers that support the feature.
- Enable QoS Pre-Classify on headend IPSec VPN routers only when both the VPN termination and QoS policies reside on the same device.

Pre-Encryption Queuing

The hardware crypto engine within a Cisco VPN router's chassis can be viewed as an internal interface that processes packets for encryption or decryption.

Before Cisco IOS Release 12.2(13)T, packets to be encrypted were handed off to the crypto engine in a *first-in-first-out* (FIFO) basis. No distinction was made between voice packets and data packets. The FIFO queuing for crypto engines is illustrated in Figure 6-11.

Figure 6-11 FIFO Crypto Engine QoS



Consider a Cisco 2651XM router deployed at a branch site configured with a full-duplex Fast Ethernet interface, a Serial E1 interface (also full duplex), and an AIM-BP encryption accelerator. The Fast Ethernet interface connects to the branch's LAN, and the serial interface connects to the Internet. These factors could limit throughput (causing a bottleneck) in this scenario:

- The clock rate of the slowest interface (in bits per second—in this case, 2 Mbps transmitted or 2 Mbps received over the E1 interface)
- The packet-forwarding rate of the router's main CPU (in packets per second)
- The crypto engine encryption/decryption rate (in packets per second)

The performance characteristics of these items further are influenced by the traffic mix, including the rates and sizes of the IP packets being switched through the router and the configured Cisco IOS switching path (process-switched, fast-switched, or CEF-switched).

Note In most hardware platforms, the packets-per-second capabilities of the router are more important for planning purposes than bits per second switched through the router. For example, if the average packet size of packets switched through the router increases from 128 bytes to 256 bytes, the packet-per-second capabilities of the main CPU are not necessarily cut in half.

The control plane requirements also factor into the CPU's utilization. These requirements are determined by the routing protocol(s) in use, the number of routes in the routing table, the overall network stability, and any redistribution requirements. Management requirements such as NTP and SNMP also add to the CPU tax. Additionally, Cisco IOS HA, QoS, multicast, and security features all consume CPU resources and must be taken into account.

Another factor in the equation is the ratio of packets switched through (and originated by) the router in relation to the number of packets selected by the crypto map's access list for encryption or decryption. If an IP GRE tunnel is being encrypted, this tends to be a large percentage of encrypted packets to total packets; if an IP GRE tunnel is not being encrypted, the ratio could be quite small.

Hardware crypto engines can become congested when their packet-processing capabilities are less than those of the router's main CPU and interface clock speeds. In such a case, the crypto engine becomes a bottleneck, or a congestion point. The crypto engine might be oversubscribed on either a momentary or (worse case) sustained basis. Such internal precrypto congestion could affect the quality of real-time applications, such as VoIP.

Cisco internal testing and evaluation has shown it to be extremely difficult for conditions to arise that cause hardware crypto engine congestion. In nearly all cases, the Cisco VPN router platform's main CPU is exhausted before reaching the limit of the crypto engine's packet-processing capabilities.

Nevertheless, Cisco provides a solution to this potential case of congestion in the rare event that a hardware crypto engine is overwhelmed so that VoIP quality will be preserved. This feature, low-latency queuing (LLQ) for IPSec encryption engines, was introduced in Cisco IOS Release 12.2(13)T.

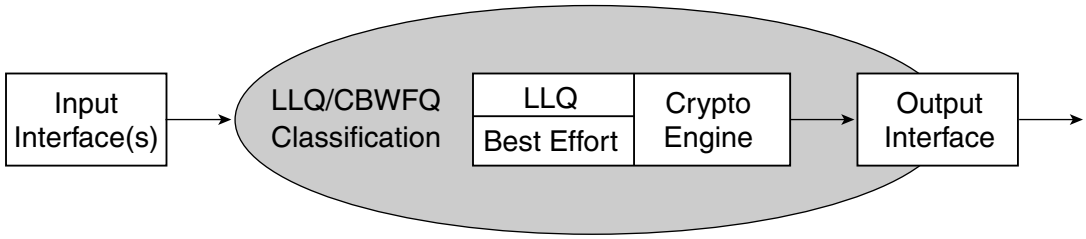
The LLQ for Crypto Engine feature provides a dual-input queuing strategy for packets being sent to the crypto engine:

- A priority or low-latency queue
- A best-effort queue

This feature is targeted at alleviating any effects of momentary or sustained oversubscription of the hardware crypto engines that could result in priority traffic (such as voice and video) experiencing quality issues. This feature is illustrated in Figure 6-12.

Note Because software-based crypto adds unacceptable latency and jitter, there are no plans to incorporate this feature for software crypto. Hardware accelerators for IPSec encryption are highly recommended.

Figure 6-12 LLQ (Dual-FIFO) Crypto Engine QoS



The classification component to segregate traffic between the priority (LLQ) queue and the best-effort queue is based on the MQC service policy on the output interface(s).

No additional configuration is required to enable LLQ for crypto engines; it is enabled internally by the presence of a service policy with an LLQ **priority** command that is applied to an output interface of an IPSec VPN router.

Traffic specified in the service policy to be assigned to the interface's priority queue (LLQ) automatically is sent to the crypto engine's LLQ. Traffic included in any CBWFQ bandwidth classes (including the default class) automatically is assigned to the crypto engine's best-effort queue.

It is possible to configure different service policies, each with different traffic assigned to an LLQ, on different interfaces. For example, perhaps voice is assigned to the LLQ of Serial1/0 and video is assigned to the LLQ of an ATM PVC. Assuming that both voice and video are to be encrypted, the question arises, which type of traffic (voice or video) will be assigned to the crypto engine's LLQ?

Because the crypto engine acts like a single interface inside the VPN router, encrypting and decrypting all outbound and inbound traffic streams for each interface on which crypto is applied, in the case of multiple service policies (on different interfaces) the crypto engine maps *all* interface priority queues (LLQ) to its LLQ and all other queues to its best-effort queue. Therefore, *both* voice and video would be assigned to the crypto engine's LLQ.

In short, the LLQ for Crypto Engine feature ensures that if packets are dropped by momentary or sustained congestion of the crypto engine, the dropped packets will be of appropriately lower priority (not VoIP packets).

Although the feature is enabled by the presence of a service policy with an LLQ **priority** statement, as with interface queuing itself, crypto-engine queuing does not actually engage prioritization through the dual-FIFO queuing strategy until the crypto engine itself experiences congestion.

The LLQ for Crypto Engine feature in Cisco IOS Software is not a prerequisite for deploying QoS for IPSec VPN implementations in a high-quality manner. As indicated previously, internal Cisco evaluations have found it extremely difficult to produce network traffic conditions that resulted in VoIP quality suffering because of congestion of the hardware crypto engine.

In general, the LLQ for Crypto Engine feature offers the most benefit under one of the following conditions:

- When implementing Cisco IOS VPN router platforms that have a relatively high amount of main CPU resources relative to crypto engine resources (these vary depending on the factors outlined earlier in this discussion).
- When the network experiences a periodic or sustained burst of large packets (for example, for video applications).

To summarize, high-quality IPSec VPN deployments are possible today without the LLQ for Crypto Engine feature in Cisco IOS software. The addition of this feature in Cisco IOS software further ensures that high-priority applications, such as voice and video, can operate in a high-quality manner even under harsh network conditions.

Anti-Replay Implications

IPSec offers inherent message-integrity mechanisms to provide a means to identify whether an individual packet is being replayed by an interceptor or hacker. This concept is called connectionless integrity. IPSec also provides for a partial sequence integrity, preventing the arrival of duplicate packets. These concepts are outlined in RFC 2401, "Security Architecture for the Internet Protocol."

When ESP authentication (**esp-sha-hmac**) is configured in an IPSec transform set, for each security association, the receiving IPSec peer verifies that packets are received only once. Because two IPSec peers can send millions of packets, a 64-packet sliding window is implemented to bound the amount of memory required to tally the receipt of a peer's packets. Packets can arrive out of order, but they must be received within the scope of the window to be accepted. If they arrive too late (outside the window), they are dropped.

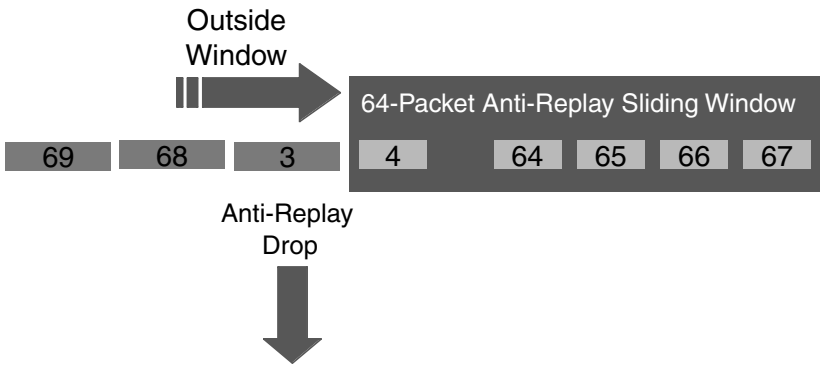
The operation of the Anti-Replay window protocol is as follows:

1. The sender assigns a unique sequence number (per security association) to encrypted packets.
2. The receiver maintains a 64-packet sliding window, the right edge of which includes the highest sequence number received. In addition, a Boolean variable is maintained to indicate whether each packet in the current window was received.
3. The receiver evaluates the received packet's sequence number:
   - If a received packet's sequence number falls within the window and was not received previously, the packet is accepted and marked as received.
   - If the received packet's sequence number falls within the window and previously was received, the packet is dropped and the replay error counter is incremented.
   - If the received packet's sequence number is greater than the highest sequence in the window, the packet is accepted and marked as received, and the sliding window is moved "to the right."
   - If the received packet's sequence number is less than the lowest sequence in the window, the packet is dropped and the replay error counter is incremented.

In a converged IPSec VPN implementation with QoS enabled, lower-priority packets are delayed so that higher-priority packets receive preferential treatment. This has the unfortunate side effect of reordering the packets to be out of sequence from an IPSec Anti-Replay sequence number perspective. Therefore, there is a concern that through the normal QoS prioritization process, the receiver might drop packets as Anti-Replay errors, when, in fact, they are legitimately sent or received packets.

Figure 6-13 provides a visualization of the process. In this example, voice packets 4 through 67 have been received, and data packet 3 was delayed and transmitted following voice packet 68. When the Anti-Replay logic is called to process packet 3, it is dropped because it is outside the left edge of the sliding window. Packets can be received out of order, but they must fall within the window to be accepted.

Figure 6-13 Anti-Replay Operation



Anti-Replay drops can be eliminated in a pure IPSec tunnel design (no encrypted IP GRE tunnel) by creating separate security associations for voice and data; voice and data packets must match a separate line in the access list referenced by the crypto map. This is implemented easily if the IP phones are addressed by network addresses (such as private RFC 1918 addresses) separate from the workstations.

However, if IPSec tunnel mode (with an encrypted IP GRE tunnel) is used for a converged network of voice and data, Anti-Replay drops impact data packets instead of voice packets because the QoS policies prioritize voice over data.

Consider the effect of packet loss on a TCP-based application: TCP is connection oriented and incorporates a flow-control mechanism within the protocol. The TCP application cannot see why a packet was dropped. A packet dropped by a service policy on a congested output interface is no different to the application than a packet lost by an Anti-Replay drop. From a *network perspective*, however, it would be more efficient to drop the packet *before* sending it over the WAN link (where bandwidth is the most expensive, only to have it dropped by the Anti-Replay mechanism on the receiving IPSec VPN router), but the location or nature of the packet loss is immaterial to the TCP driver.

Anti-Replay drops of data traffic flows are usually in the order of 1 percent to 1.5 percent on IPSec VPN links that experience sustained congestion and have queuing engaged (without any additional policy tuning).

Output drops on the output WAN interface, however, tend to be few, if any, and certainly are far fewer than those dropped by Anti-Replay. This is because Anti-Replay triggers packet drops more aggressively than the output service policy, which is a function of the size of the output queues and the number of defined classes.

By default, each CBWFQ class receives a queue with a length of 64 packets. This can be verified with the **show policy interface** verification command. Meanwhile, the receiving IPSec peer has a *single* 64-packet Anti-Replay window (per IPSec Security Association) with which to process all packets from all LLQ and CBWFQ bandwidth classes.

So, it stands to reason that the Anti-Replay mechanism on the receiving VPN router will be more aggressive at dropping packets delayed by QoS mechanisms preferential to VoIP than the service policy at the sending router. This is because of the size mismatch of the queue depth on the sender's output interface (multiple queues of 64 packets each) compared to the width of the receiver's Anti-Replay window (a single sliding window of 64 packets per SA). As more bandwidth classes are defined in the policy map, this mismatch increases. As mentioned, this is an inefficient use of expensive WAN/VPN bandwidth because packets are transmitted only to be dropped before decryption.

The default value of 64 packets per CBWFQ queue is designed to absorb bursts of data traffic and delay rather than drop those packets. This is optimal behavior in a non-IPSec–enabled network.

When IPSec authentication is configured (**esp-sha-hmac**) in the network, the scenario can be improved by reducing the queue limit (max threshold) of the bandwidth classes of the sender's output service policy so that it becomes more aggressive at dropping packets than buffering or delaying them. Extensive lab testing has shown that such queue-limit tuning can reduce the number of Anti-Replay drops from 1 percent to less than a tenth of percent (< 0.1 percent). This is because decreasing the service policy queue limits causes the sender's output service policy to become more aggressive in dropping instead of significantly delaying packets (which occurs with large queue depths). This, in turn, decreases the number of Anti-Replay drops.

As a rule of thumb, the queue limits should be reduced in descending order of application priority. For example, the queue limit for Scavenger should be set lower than the queue limit for a Transactional Data class, as is shown later in Example 6-3 through Example 6-6.

Note In many networks, the default queue limits and IPSec Anti-Replay performance are considered acceptable. A modification of queue-limit values entails side effects on the QoS service policy and related CPU performance. When queue limits are tuned, a careful eye should be kept on CPU levels.

Note As of IOS 12.3(14)T another potential remedy to QoS/IPSec Anti-Replay issues became available with the ability to expand or disable the IPSec Anti-Replay window. However it should be kept in mind that the IETF IPSec standards define the Anti-Replay window-sizes to be 64 packets and as such, this would not be a standards-compliant solution. On the other hand, the presented solution of tuning of the QoS queue-limits is fully standards-compliant.

The documentation for IOS feature to allow the expanding or disabling of IPSec Anti-Replay is at: http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123newft/123t/123t_14/gt_iarwe.htm

Control Plane Provisioning

As discussed in Chapter 3, "WAN Aggregator QoS Design,", Cisco IOS Software has an internal mechanism for protecting control traffic, such as routing updates, called PAK_priority.

PAK_priority marks routing protocols to DSCP CS6, but it currently does not protect IPSec control protocols, such as ISAKMP (UDP port 500). Therefore, it is recommended to provision an explicit CBWFQ bandwidth class for control plane traffic, including ISAKMP, as shown in Example 6-2.

Example 6-2 Protecting IPSec Control Plan Traffic

```
!
class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6
  match access-group name IKE                ! References ISAKMP ACL
!
!
policy-map V3PN
 ...
  class INTERNETWORK-CONTROL
   bandwidth percent 5                       ! Control Plane provisioning
...
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp      ! ISAKMP ACL
```

```
!
```

Site-to-Site V3PN QoS Designs

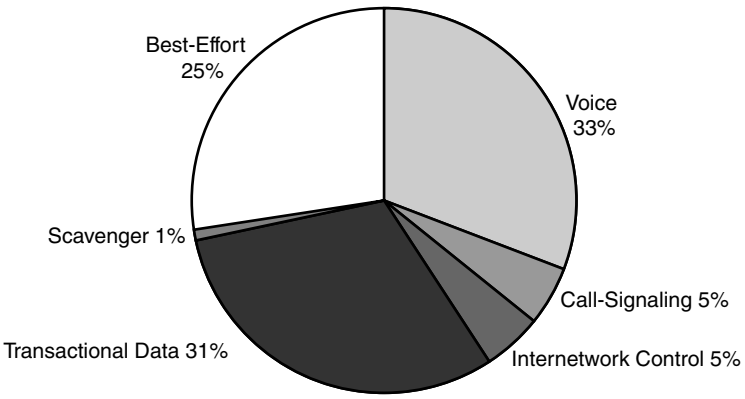
As with WAN and MPLS VPN QoS models, site-to-site V3PN QoS models can range from a basic number of classes (in this case, the minimum number of recommended classes is 6) to a complex QoS Baseline model (11 classes). Each enterprise must determine present needs and comfort level of QoS complexity, along with future needs, to more easily migrate to progressively more complex QoS models, as required.

Six-Class Site-to-Site V3PN Model

The six-class V3PN model technically should be referred to as V2PN because it includes provisioning for only voice (not for video) over an IPSec VPN. Voice is protected explicitly with LLQ; call signaling and control plane traffic also explicitly are protected through CBWFQ classes. This model includes a preferential data class (Transactional Data) and a deferential class (Scavenger, which is squelched to the minimum configurable amount: 1 percent). If the queue limits are tuned to minimize Anti-Replay drops, the queue limit for Transactional Data should be set higher than the queue limit for class default.

An example six-class V3PN model, which is suitable for link speeds up to and including T1/E1 speeds, is illustrated in Figure 6-14 and detailed in Example 6-3.

Figure 6-14 Six-Class Site-to-Site V3PN Model



Example 6-3 Six-Class Site-to-Site V3PN Model Configuration Example

```
!
class-map match-all VOICE
  match ip dscp ef                          ! VoIP
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                         ! New Call-Signaling
  match ip dscp af31                        ! Old Call-Signaling
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                         ! IP Routing
  match access-group name IKE               ! References ISAKMP ACL
 class-map match-all TRANSACTIONAL-DATA
  match ip dscp af21 af22                   ! Transactional-Data
 class-map match-all SCAVENGER
  match ip dscp cs1                         ! Scavenger
```

```
!
!
 policy-map SIX-CLASS-V3PN-EDGE
  class VOICE
    priority percent 33                          ! VoIP gets 33% LLQ
  class CALL-SIGNALING
    bandwidth percent 5                          ! Call-Signaling provisioning
  class INTERNETWORK-CONTROL
    bandwidth percent 5                          ! Control Plane provisioning
  class TRANSACTIONAL-DATA
    bandwidth percent 31                         ! Transactional-Data provisioning
    queue-limit 20                               ! Optional: Anti-Replay tuning
  class SCAVENGER
    bandwidth percent 1                          ! Scavenger class is throttled
    queue-limit 1                                ! Optional: Anti-Replay tuning
  class class-default
    bandwidth percent 25                         ! Best Effort needs BW guarantee
    queue-limit 16                               ! Optional: Anti-Replay Tuning
!
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp          ! ISAKMP ACL
!
```

Note Currently, only distributed platforms support the **queue-limit** command in conjunction with WRED commands. However, this command combination will be available on nondistributed platforms with the release of Consistent QoS Behavior, as discussed in Chapter 3, "WAN Aggregator QoS Design.".

Verification commands:

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

Eight-Class Site-to-Site V3PN Model

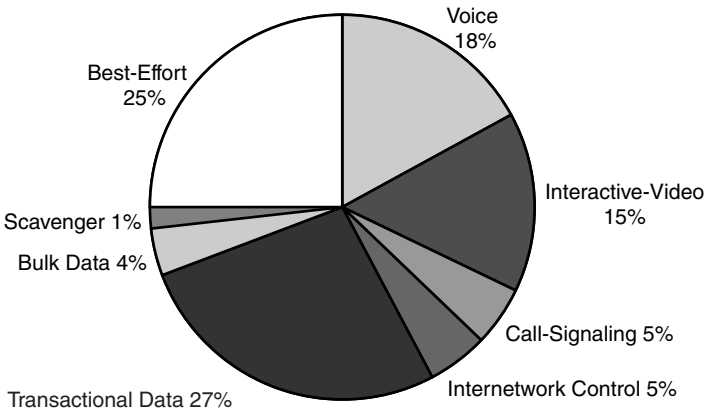
With the addition of a class for Interactive-Video, this model more accurately lives up to its V3PN name. Interactive-Video (as with VoIP) must be provisioned adequately to include IPSec encryption overhead, but (unlike VoIP) there are no clean formulas for calculating the required incremental bandwidth. This is because video packet sizes and packet rates vary significantly and are largely a function of the degree of motion within the video images being transmitted.

On an unencrypted topology, the guideline for provisioning for Interactive-Video is to overprovision the LLQ by 20 percent. Although this conservative guideline holds true in most videoconferencing scenarios, which usually consist of relatively minor motion (meetings, conferences, talking heads, and so on), in some cases this rule proves inadequate over an IPSec VPN. In such cases, the LLQ might need to be provisioned to the stream's rate plus 25 percent or more. The need to increase the bandwidth for Interactive-Video's LLQ will be apparent if drops (in excess of 1 percent) appear under this class when using the verification command **show policy interface**.

The other class added to this model is a separate class for Bulk Data applications (which will be constrained if congestion occurs, to prevent long sessions of TCP-based flows from dominating the link). Notice that the queue limit of the Bulk Data class has been reduced to below the queue limit of the Best-Effort class. This is because of the relative ranking of application priority: During periods of congestion, the Bulk Data class (large TCP-based file operations that operate mainly in the background) is prevented from dominating bandwidth away from the Best-Effort class (as a whole).

This policy is suitable for link speeds of 3 Mbps or higher. However, in an IPSec environment, it is good to remember that load-sharing GRE tunnels over multiple physical interfaces exacerbate Anti-Replay drops. Whenever possible, a single physical interface should be used to achieve these higher speeds. Another consideration to bear in mind is that higher-end platforms, such as the 2691 and 3700- or 7200-series routers, are required to perform crypto at higher speeds. The Eight-Class Site-to-Site V3PN model is illustrated in Figure 6-15 and detailed in Example 6-4.

Figure 6-15 Eight-Class Site-to-Site V3PN Model



Example 6-4 Eight-Class Site-to-Site V3PN Model Configuration Example

```
!
 class-map match-all VOICE
  match ip dscp ef                          ! VoIP
 class-map match-all INTERACTIVE-VIDEO
  match ip dscp af41 af42                   ! Interactive-Video
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                         ! Old Call-Signaling
  match ip dscp af31                        ! New Call-Signaling
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                         ! IP Routing
  match access-group name IKE               ! References ISAKMP ACL
 class-map match-all TRANSACTIONAL-DATA
  match ip dscp af21 af22                   ! Transactional-Data
 class-map match-all BULK-DATA
  match ip dscp af11 af12                   ! Bulk Data
 class-map match-all SCAVENGER
  match ip dscp cs1                         ! Scavenger
!
 policy-map EIGHT-CLASS-V3PN-EDGE
  class VOICE
    priority percent 18               ! VoIP gets 18% LLQ
  class INTERACTIVE-VIDEO
    priority percent 15               ! IP/VC gets 15% LLQ
class CALL-SIGNALING
    bandwidth percent 5               ! Call-Signaling provisioning
 class INTERNETWORK-CONTROL
    bandwidth percent 5               ! Control Plane provisioning
  class TRANSACTIONAL-DATA
    bandwidth percent 27              ! Transactional-Data provisioning
    queue-limit 18                    ! Optional: Anti-Replay tuning
  class BULK-DATA
    bandwidth percent 4               ! Bulk-Data provisioning
    queue-limit 3                     ! Optional: Anti-Replay tuning
```

```
  class SCAVENGER
    bandwidth percent 1               ! Scavenger class is throttled
    queue-limit 1                     ! Optional: Anti-Replay tuning
  class class-default
    bandwidth percent 25              ! Best Effort needs BW guarantee
    queue-limit 16                    ! Optional: Anti-Replay Tuning
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp    ! ISAKMP ACL
!
```

Verification commands:

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

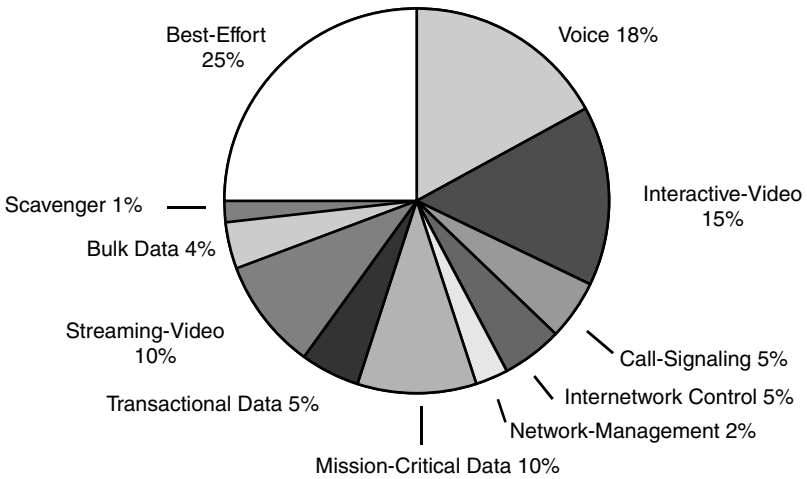
QoS Baseline (11-Class) Site-to-Site V3PN Model

Building on the previous model, three new classes are added: Network-Management, Mission-Critical Data, and Streaming-Video.

This model also is suitable only for high-speed (3 Mbps and above) links, with the same caveats regarding multiple physical links aggravating Anti-Replay drops and the requirement of using newer platforms to perform crypto at these speeds. The queue limits have been tuned to reflect relative application priority.

The QoS Baseline V3PN model, suitable for 3-Mbps link speeds and higher, is illustrated in Figure 6-16 and detailed in Example 6-5.

Figure 6-16 QoS Baseline (Eleven-Class) Site-to-Site V3PN Model



Example 6-5 QoS Baseline (Eleven-Class) Site-to-Site V3PN Model Configuration Example

```
!
 class-map match-all VOICE
  match ip dscp ef                          ! VoIP
 class-map match-all INTERACTIVE-VIDEO
  match ip dscp af41 af42                   ! Interactive-Video
 class-map match-any CALL-SIGNALING
```

```
  match ip dscp cs3                          ! Old Call-Signaling
  match ip dscp af31                         ! New Call-Signaling
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                          ! IP Routing
  match access-group name IKE                ! References ISAKMP ACL
 class-map match-all NET-MGMT
  match ip dscp cs2                          ! Network Management
 class-map match-all MISSION-CRITICAL-DATA
  match ip dscp 25                           ! Interim MC Data
 class-map match-all TRANSACTIONAL-DATA
  match ip dscp af21 af22                    ! Transactional Data
 class-map match-all STREAMING-VIDEO
  match ip dscp cs4                          ! Streaming Video
 class-map match-all BULK-DATA
  match ip dscp af11 af12                    ! Bulk Data
 class-map match-all SCAVENGER
  match ip dscp cs1                          ! Scavenger
!
!
policy-map QOSBASELINE-V3PN-EDGE
  class VOICE
    priority percent 18              ! VoIP gets 18% LLQ
  class INTERACTIVE-VIDEO
    priority percent 15              ! IP/VC gets 15% LLQ
  class CALL-SIGNALING
    bandwidth percent 5              ! Call-Signaling provisioning
  class INTERNETWORK-CONTROL
    bandwidth percent 5              ! Control Plane provisioning
  class NET-MGMT
    bandwidth percent 2              ! Network Management provisioning
  class MISSION-CRITICAL-DATA
    bandwidth percent 10             ! Mission-Critical Data provisioning
    queue-limit 6                    ! Optional: Anti-Replay tuning
  class TRANSACTIONAL-DATA
    bandwidth percent 5              ! Transactional-Data provisioning
    queue-limit 3                    ! Optional: Anti-Replay tuning
  class STREAMING-VIDEO
    bandwidth percent 10             ! Streaming-Video provisioning
    queue-limit 6                    ! Optional: Anti-Replay tuning
  class BULK-DATA
    bandwidth percent 4              ! Bulk-Data provisioning
    queue-limit 3                    ! Optional: Anti-Replay tuning
  class SCAVENGER
    bandwidth percent 1              ! Scavenger throttling
    queue-limit 1                    ! Optional: Anti-Replay tuning
class class-default
    bandwidth percent 25             ! Best Effort needs BW guarantee
    queue-limit 16                   ! Optional: Anti-Replay tuning
!
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp     ! ISAKMP ACL
!
```

Verification commands:

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

At remote sites, these policies can be applied to the physical interfaces connecting them to the service provider (provided that the SLAs are for line rates—otherwise, shapers must be used). For example, if the service provider guarantees a full T1 rate to a remote site and the access medium is Frame Relay, the

service policy can be applied directly to the main Frame Relay interface. If the service provider guarantees only £ 768 kbps across the same link, Frame Relay traffic shaping (either legacy or class-based FRTS) combined with FRF.12 must be used at the remote site.

For the central site(s) WAN aggregators, however, unique considerations exist. These are discussed in the next section.

Headend VPN Edge QoS Options for Site-to-Site V3PNs

IPSec V3PNs can be configured in various ways at the central sites. Some enterprises simply overlay V3PNs on top of their existing private WANs; others subscribe to service providers that offer classes of service within their clouds. Many enterprises deploy VPN headends behind WAN aggregation routers to distribute CPU loads, while some perform encryption and QoS on the same box. Each of these options presents considerations on how V3PN policies optimally are applied on WAN aggregation routers.

For enterprises that have overlaid IPSec VPNs on top of their private WAN topologies, the V3PN policies should be applied to the leased lines or Frame Relay/ATM PVCs, as described in Chapter 3, "WAN Aggregator QoS Design."

For enterprises that are subscribing to service providers that offer PE-to-CE QoS classes (including enterprises that are deploying IPSec VPNs *over* MPLS VPNs), V3PN policies need to be applied on the CE-to-PE links (complete with any re-marking that the service provider requires to map into these service provider classes), as described in Chapter 5, "MPLS VPN QoS Design."

For enterprises that are subscribing to service providers that do not offer explicit QoS (beyond an SLA) within their cloud and are using VPN headends behind WAN aggregation routers, the V3PN service policies would be applied to the WAN aggregator's (CE-to-PE) physical links. This prioritizes packets (by applications) and relies on the service provider's SLA to ensure that the delivery largely reflects the priority of the packets as they are handed off to the service provider. Such a configuration would require only a single QoS policy for a WAN aggregator (albeit, on a high-speed interface), but at the same time, it would involve an increased dependence on the service provider's SLA to deliver the desired QoS service levels.

When the VPN headend routers have adequate CPU cycles to perform QoS, another option exists: hierarchical MQC policies that shape and queue (within the shaped rate) and are applied on a per-tunnel basis. A sample of per-tunnel hierarchal QoS policies is shown in Example 6-6. As with previous shaping design recommendations, the shaper is configured to shape to 95 percent of the remote site's line rate.

Note It is critical to keep an eye on CPU levels when IPSec VPN encryption *and* per-tunnel QoS policies are applied on the same router. CPU levels, in general, should not exceed 75 percent during normal operating conditions. Configuring hierarchical shaping and queuing policies on a per-tunnel (per-SA) basis to a large number of sites could be very CPU intensive, especially when such sites experience periods of sustained congestion.

Example 6-6 Per-Tunnel Hierarchical Shaping and Queuing MQC Policy for VPN Headends/WAN Aggregators

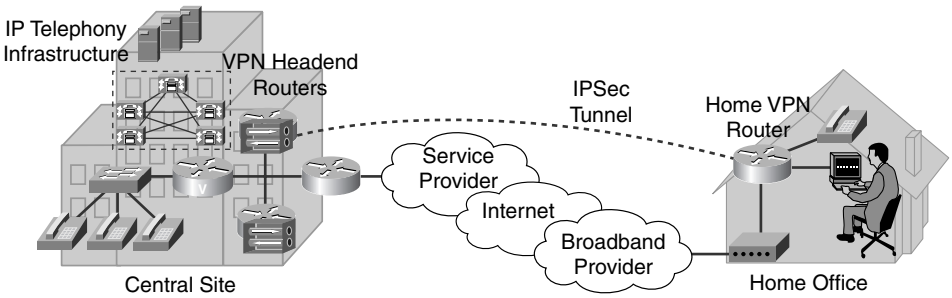
```
!
policy-map SHAPING-T1-TUNNEL
  class class-default
   shape average 1460000 14600 0       ! Shaped to 95% of T1 line-rate
   service-policy V3PN-EDGE            ! Nested queuing policy
!
```

```
!
interface Tunnel120
 description VPN Pipe to V3PN Site#120 (T1 Link)
 bandwidth 1536
 ip address 10.10.120.1 255.255.255.252
 ip mtu 1420
 service-policy output SHAPING-T1-TUNNEL     ! Policy applied to tunnel int
 qos pre-classify                            ! Performance recommandation
 tunnel source 192.168.1.1
 tunnel destination 192.168.2.2
 crypto map VPN
!
```

Teleworker V3PN QoS Considerations

Organizations constantly are striving to reduce costs and improve productivity and employee retention. Teleworker solutions address these organizational objectives by giving employees the ability to work from home with compatible quality, functionality, performance, convenience, and security, as compared to working in an office location. Teleworker solutions that provide such functionality have been branded "enterprise class" by Cisco Marketing and are illustrated in Figure 6-17.

Figure 6-17 Enterprise Teleworker Design



Telecommuter solutions include these main benefits:

- **Increased productivity**—On average, employees spend 60 percent of their time or less at their desks, yet this is where the bulk of investment is made in providing access to corporate applications. Providing similar services at an employee's residence, for a relatively minor investment, significantly can increase productivity gains.
- **Business resilience**—Employees can be displaced from their normal workplace by natural events (such as winter storms, hurricanes, or earthquakes), health alerts (such as SARS), man-made events (such as travel restrictions or traffic conditions), or simply family-related events, such as sick children or home repairs. These disruptions significantly can impact an organization's processes. Providing employees with central site–equivalent access to applications and services in geographically dispersed locations (such as home offices) creates a built-in back-up plan to keep business processes functioning in unforeseen circumstances.
- **Cost savings**—A traditional remote worker setup involves toll charges for dial-up and additional phone lines. Integrating services into a single, broadband-based connection can eliminate these charges while delivering superior overall connectivity performance.

- **Security**—Demands for access to enterprise applications outside the campus are stretching the limits of security policies. Teleworking over IPSec VPNs offers inherent security provided by encryption of all traffic, including data, voice, and video. Also critical is integrating firewall and intrusion-detection capabilities, along with finding ways to easily accommodate both corporate and personal users who share a single broadband connection (the "spouse-and-children" concern, which will be discussed shortly).
- **Employee recruitment and retention**—In the past, enterprises recruited employees in the locations where corporate offices were located. Today enterprises need the flexibility of hiring skilled employees wherever the skills exist and need to integrate remote workers into geographically dispersed teams with access to equivalent corporate applications.

Although QoS designs for IPSec V3PN teleworker scenarios share many of the same concerns of site-to-site V3PN scenarios, additional specific considerations relating to teleworker deployment models and broadband access technologies need to be taken into account.

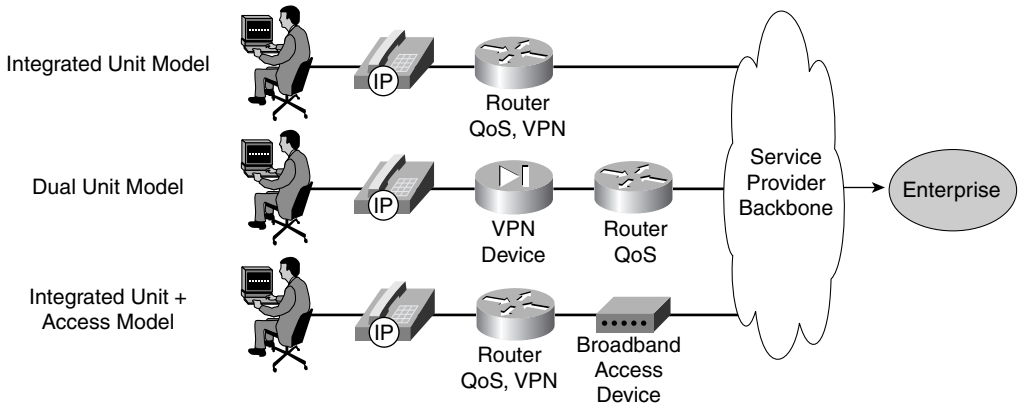
Teleworker Deployment Models

Enterprise teleworker deployment models must provide the following services:

- **Basic services**—These include NAT, DHCP, IP routing, and multiple Ethernet connections for home office devices and broadband connection (attachment to the WAN circuit cable, DSL, ISDN, or wireless network).
- **QoS services**—These include classification, prioritization, and, in some cases, shaping of traffic.
- **VPN/security services**—These include encryption of traffic to the main site and firewall functionality for the home office.

Given these requirements, there are three main deployment models for Enterprise teleworker V3PN solutions, as illustrated in Figure 6-18. These include the Integrated Unit Model, the Dual-Unit Model, and the Integrated Unit + Access Model.

Figure 6-18 Enterprise Teleworker Deployment Models



Integrated Unit Model

In the Integrated Unit Model, a single device (such as a Cisco 837 or 1700-series router) provides basic services, QoS services, and VPN/security services. Furthermore, the device connects directly to the broadband media.

Note The Cisco routers used in these scenarios require an IP/FW/PLUS 3DES Cisco IOS Software feature set. This feature set would include support for LLQ/CBWFQ with class-based hierarchical traffic shaping, and also support for Easy VPN (EZVPN), PKI, IDS, AES, URL filtering, and EIGRP.

Advantages include the following:

- Single-device deployment and management
- Adaptability for service provider fully managed services (transport, QoS, IP Telephony application)
- Potential cost savings

Disadvantages include the following:

- Availability of a single device at an appropriate cost with the features and performance required
- No single unit for some broadband access circuit types

The Integrated Model is a preferred model for service providers offering a fully managed V3PN teleworker service.

From a QoS design perspective, this model is highly similar to a site-to-site V3PN, except that it interfaces with a broadband media instead of a traditional WAN access media.

Dual-Unit Model

In this model, one device (such as a Cisco 831, 837, or 1700-series router) provides basic services and QoS, and a second device (a Cisco router or a PIX 501 or even a VPN 3002) performs security/VPN services.

Advantages include the following:

- **Granularity of managed services**—Service providers can manage broadband access while enterprises manage VPN/security and private network addressing because these are two different units.
- **Media independence**—Because the VPN/security device is separate from the router connecting to the broadband circuit, the same VPN device can be used for cable, DSL, wireless, and ISDN by changing the router model or module in the router. This is especially valuable if one enterprise must support teleworkers with different broadband circuit types (such as DSL, cable, and ISDN).

Disadvantages include the following:

- Two units packaged for deployment
- Ongoing management of two devices
- The cost for two devices

From a QoS design perspective, this model is no different from the previous (Integrated Unit) model.

Integrated Unit + Access Model

In this third model, a single router (such as a Cisco 831 or 1700-series router) provides basic services, QoS services, and VPN/security services. However, the router does not connect directly to the broadband access media; it connects (through Ethernet) to a broadband access device (such as a DSL or cable modem).

Advantages include the following:

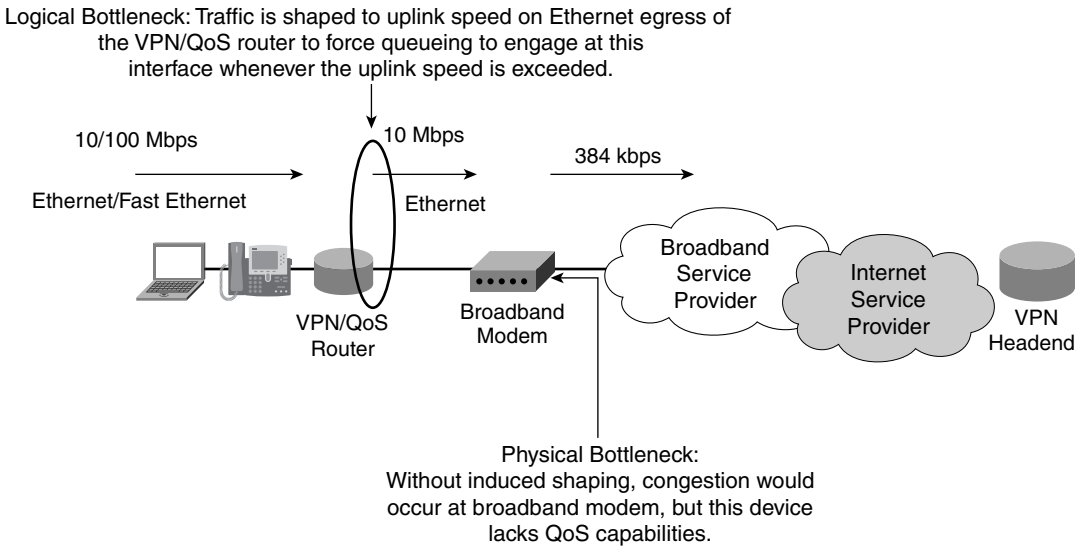
- Cost savings realized by using existing broadband-access devices

- Simplified provisioning because one size fits all, regardless of broadband access media
- Solution support even when no router interface for a specific broadband circuit type is available

Disadvantages include the following:

- Increased troubleshooting complexity because most broadband-access devices (modem) are not intelligent and, therefore, cannot be queried, managed, or controlled.
- Additional QoS complexity because, in this model, the router does not control the broadband circuit and, thus, must perform hierarchical shaping, as shown in Figure 6-19.

Figure 6-19 Hierarchical QoS Requirements for Integrated Unit + Access Teleworker Deployment Model



Because of the media-specific encapsulation overhead requirements (discussed in the following sections), it is recommended to shape to 95 percent of broadband link for cable and 70 percent of the uplink rate for DSL.

A hierarchical shaping policy that forces queuing to engage for a 384-kbps cable broadband connection is shown in Example 6-7.

Example 6-7 Hierarchical Shaping and Queuing MQC Policy for a 384-kbps Cable Connection

```
!
policy-map SHAPE-384-CABLE
 class class-default
   shape average 364800 3640          ! Shapes to 95% of 384 kbps cable link
   service-policy V3PN-TELEWORKER     ! Nested V3PN Teleworker queuing policy
!
...
!
interface Ethernet0
 service-policy output SHAPE-384-CABLE     ! Shaper applied to LAN interface
!
```

The Integrated Unit + Access Model is a preferred model for enterprise V3PN teleworker deployments because it completely abstracts any service provider– or access media–specific requirements (a one-size-fits-all solution).

Broadband-Access Technologies

In North America, there are four main broadband-access technology options for telecommuter scenarios: DSL, cable, ISDN, and wireless. DSL and cable are by far the dominant broadband technologies.

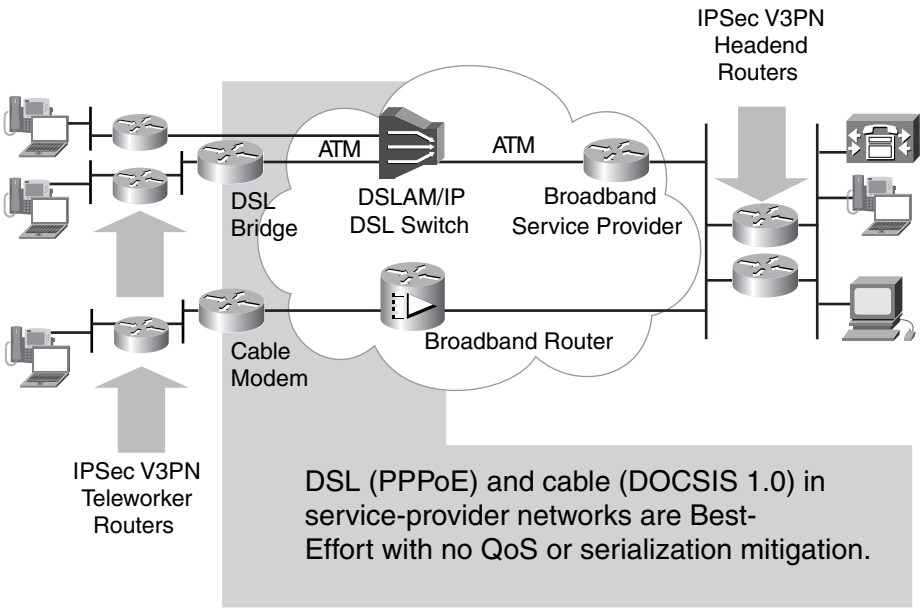
Because of per-minute costs, ISDN is used considerably less; however, ISDN flat rate is becoming available and will make ISDN a good option for areas where DSL or cable is not available. Last-mile wireless is a new option that is being piloted in certain areas to determine viability.

The minimum recommended broadband data rate for most deployments is 160 kbps (uplink)/860 kbps (downlink). Data rates below this speed require more troubleshooting by support staff and are less likely to provide acceptable voice quality. The recommended data rate for V3PN teleworker deployments is 256 kbps (uplink)/1.4 Mbps (downlink) or higher rates. Although V3PN can be deployed at rates less than 160 kbps/860 kbps, generally the voice quality at that service level is in the cell phone quality range, and support costs are higher.

Because QoS design for ISDN was discussed in Chapter 3, "WAN Aggregator QoS Design,", and because wireless as a last-mile broadband technology has yet to gain wide deployment, this section focuses only on DSL and cable broadband technologies.

DSL and cable topologies are illustrated in Figure 6-20. Cable is a shared medium between the teleworker's cable modem and the broadband provider's cable headend router. DSL is a dedicated circuit between the teleworker's DSL modem (bridge) and the DSL Access Multiplexer (DSLAM). Both cable and DSL offerings utilize shared uplinks between these aggregation devices and the service provider's core network. QoS typically is not provisioned within the broadband service provider's cloud in either medium. This lack of QoS within the broadband cloud underscores the requirement of adequate QoS provisioning at the endpoints of the VPN tunnels.

Figure 6-20 DSL and Cable Topologies



Digital Subscriber Line

DSL service features a dedicated access circuit and offers a service similar to Frame Relay or Asynchronous Transfer Mode (ATM), in which a single permanent virtual circuit (PVC) is provisioned from the home office to the service provider aggregation point. DSL has a variety of speeds and encoding schemes.

Most service providers today offer residential Asynchronous Digital Subscriber Line (ADSL). ADSL provides for asymmetric speeds (with the downstream rate greater than the upstream rate). Asymmetrical links are a better choice and benefit for the telecommuter because the greater the downlink bandwidth is, the less the need is for QoS. (Slower-speed uplinks slow TCP sender transmission rates to more manageable levels.) Because QoS is not generally available by broadband service providers, such uplink/downlink speed mismatches are desirable.

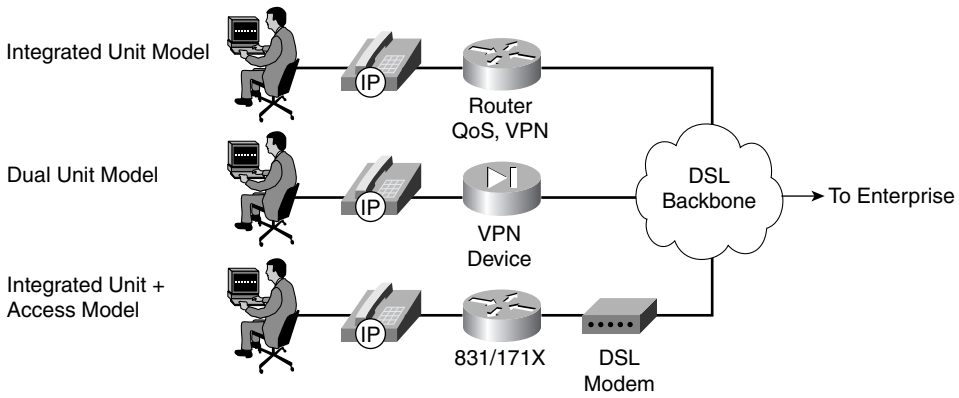
Residential access speeds are generally 128 to 384 kbps upstream and 608 kbps to 1.5 Mbps downstream. For ADSL circuits with upstream speeds less than 256 kbps, G.729 VoIP codecs are recommended.

ADSL also utilizes a single best-effort PVC using RFC 2516–based PPP over Ethernet (PPPoE) encapsulation. In DSL networks, delay and jitter are very low but are not guaranteed. Because PPPoE is used, no LFI is available in service provider DSL networks. In the future, QoS at Layer 2 might be available across service provider networks using familiar ATM variable bit-rate (VBR) definitions.

Single-pair high bit-rate DSL (G.SHDSL) is the new high-speed standard for business DSL. Most residences will continue to be served by ADSL, while small business and branch offices likely will use G.SHDSL. G.SHDSL offers varying rates controlled by the service provider; the upstream and downstream rates are the same speed (symmetric). G.SHDSL is seen as an eventual replacement for T1s in the United States and will become increasingly available from more service providers.

The telecommuter deployment options for DSL circuits include all three teleworker deployment models: Integrated Unit, Dual-Unit, and Integrated Unit + Access, as shown in Figure 6-21.

Figure 6-21 Teleworker Deployment Models for DSL



Cable

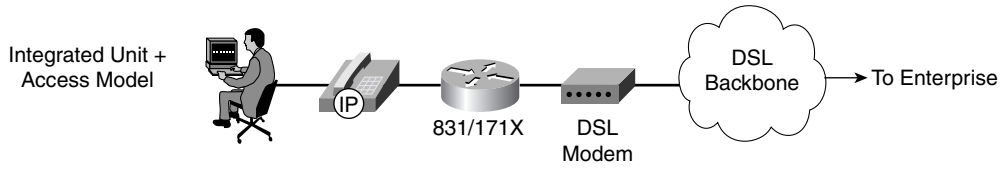
Cable offers a shared service with symmetric speeds varying from 100 kbps to 4 Mbps. In the past, delay and jitter varied too greatly over cable, making it unsuitable for VoIP.

The common installed base of cable services today is made up of Data-over-Cable Service Interface Specifications (DOCSIS) 1.0. No LFI is available with DOCSIS 1.0, although DOCSIS 1.1 defines fragmentation and interleaving mechanisms.

DOCSIS 1.1 also provides the capabilities to shape traffic at Layer 2 before transmission over the cable network. Although the circuit and frequencies physically are shared, access to the medium can be controlled by the headend so that a device can be guaranteed specified bandwidth.

At the time of this writing, the only recommended deployment model for cable is the Integrated Unit + Access Model, as shown in Figure 6-22.

Figure 6-22 Teleworker Deployment Model for Cable



Bandwidth Provisioning

A few key differences with respect to bandwidth provisioning need to be taken into account for broadband teleworker scenarios, as compared to site-to-site VPNs.

The first is that usually only a single call needs to be provisioned for (unless two teleworkers share a single residential broadband connection, which is rarely the case). If bandwidth is low, G.729 codecs are recommended. The second key bandwidth-provisioning consideration is the inclusion of the overhead required by the broadband access technologies, whether DSL or cable.

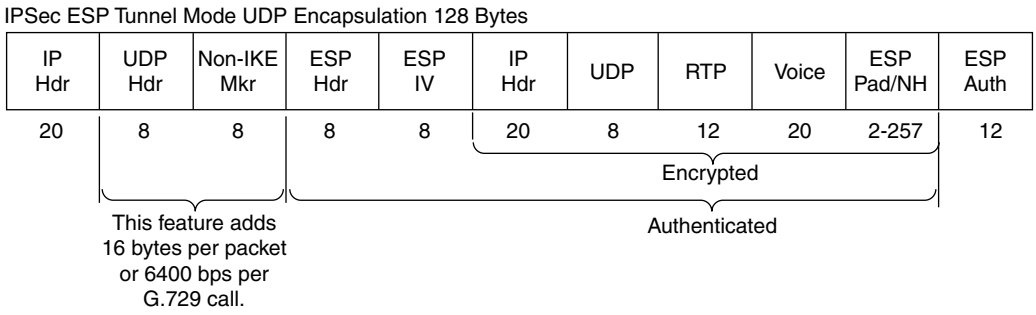
Note Sometimes multicast support is not required in a teleworker environment. For example, routing protocols (which depend on multicast support) might not be required by teleworkers (because default gateways are usually sufficient in this context). In such cases, IPSec tunnel mode (no encrypted IP GRE tunnel) can be used to achieve additional bandwidth savings of 24 bytes per packet.

For the remainder of this chapter, IPSec tunnel mode (no encrypted IP GRE tunnel) is the assumed mode of operation.

NAT Transparency Feature Overhead

Beginning in Cisco IOS Release 12.2(13)T, NAT transparency was introduced and is enabled by default, provided that both peers support the feature. It is negotiated in the IKE exchange between the peers. This feature addresses the environment in which IPSec packets must pass through a NAT/pNAT device, and it adds 16 bytes to each voice and data packet. The overhead that this feature adds is shown in Figure 6-23.

Figure 6-23 NAT Transparency Feature (Layer 3) Overhead



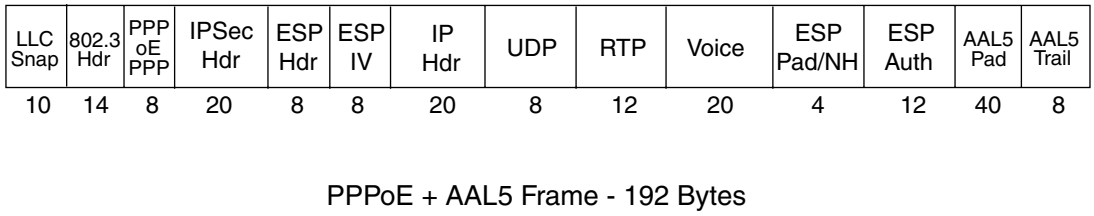
No additional overhead is caused by NAT transparency on a G.729 call over DSL (PPPoE/AAL5) because there is enough AAL5 cell padding to absorb the 16 additional bytes per packet (discussed in more detail in the following section). In cable implementations, these additional 16 bytes increase the number of bits on the wire and, thus, the overall bandwidth consumption. At the headend, NAT transparency increases the bandwidth consumption of VoIP traffic by 1 Mbps for approximately every 82 concurrent G.729 calls; on the teleworker router, NAT transparency increases the bandwidth required by 6.4 kbps per G.729 call.

Unless there is a need to implement this feature, the recommendation is to disable it when bandwidth conservation is a requirement. This feature can be disabled with the **no crypto ipsec nat-transparency udp-encapsulation** global configuration command.

DSL (AAL5 + PPPoE) Overhead

For DSL broadband connections, PPPoE is the most commonly implemented deployment. Given the 112-byte, Layer 3 size of an IPSec (only) encrypted G.729 VoIP call, 40 bytes of PPP, PPPoE, Ethernet, and ATM AAL5 headers and trailers are added. Additionally, the pre-ATM Software Segmentation and Reassembly (SAR) engine is required to pad the resulting 152-byte packet to the nearest multiple of 48 (each fixed-length ATM cell can transport only a 48-byte payload). Figure 6-24 shows the resulting 192-byte pre-ATM packet.

Figure 6-24 IPSec-Encrypted G.729 Packet Size Through AAL5 + PPPoE Encapsulation



The 192 bytes of cell payload are incorporated through ATM SAR into the (48-byte) payloads of four 53-byte ATM cells (192 /48 = 4 cells).

Therefore, the bandwidth required for an IPSec (only) encrypted G.729 VoIP call over DSL is calculated as follows:

53 bytes per cell
· 4 cells per IPSec-encrypted VoIP packet
212 bytes per (192-byte) AAL5/PPPoE IPSec VoIP packet
· 50 packets per second

10,600 bytes per second
· 8 bits per byte
84,800 bits per second

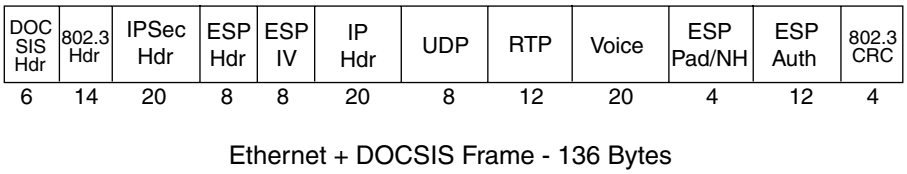
Thus, an encrypted G.729 call requires 85 kbps of bandwidth, while an encrypted G.711 requires seven ATM cells or 148,400 bps on the wire.

This results in the LLQ configuration values of 85 kbps (**priority 85**) for a G.729 VoIP call and 150 kbps for a G.711 (**priority 150**) VoIP call, respectively, for DSL.

Cable Overhead

For cable deployments, the Layer 2 overhead is less than that for DSL. The IPSec packet is encapsulated in an Ethernet header (and trailer) that includes a 6-byte DOCSIS header, as shown in Figure 6-25.

Figure 6-25 IPSec-Encrypted G.729 Packet Size Through Ethernet and DOCSIS Encapsulation



If baseline privacy is enabled (baseline privacy encrypts the payload between a cable modem and the headend), the extended header is used, adding another 5 bytes.

The packet size of a G.729 call (with a zero-length extended header) is 136 bytes or 54,400 bps (at 50 pps); a G.711 call is 280 bytes or 112,000 bps (at 50 pps).

To simplify configuration and deployment, the values of 64 kbps (for G.729) and 128 kbps (for either G.729 or G.711) can be used for priority queue definition for cable.

Asymmetric Links and Unidirectional QoS

With both DSL and cable, the uplink connection can be enabled with QoS, either in the form of a service policy on the DSL (ATM PVC) interface or through a hierarchical MQC service policy that shapes the uplink and prioritizes packets within the shaped rate on the Ethernet interface. This half of the link is under the enterprise's control and easily can be configured.

The downlink connection is under the control of the broadband service provider, and any QoS policy must be configured by the service provider; however, most service providers do not offer QoS-enabled broadband services. This is usually because DSL providers often have implemented non-Cisco equipment (DSLAM or other ATM concentration devices) that typically have few or no QoS features available. Cable providers, on the other hand, might have an option to enable QoS as DOCSIS 1.1 becomes more widely deployed.

Fortunately, most service offerings are asymmetrical. For example, consider a circuit with a 256-kbps uplink and 1.5-Mbps downlink. The downlink rarely is congested to the point of degrading voice quality.

Testing in the Cisco Enterprise Solutions Engineering labs has shown that when congestion is experienced on the uplink, the resulting delay of (TCP) data packet acknowledgments automatically decreases the arrival rate of downlink data traffic so that downlink congestion does not occur. To

summarize the results of these tests: Asymmetrical links are preferred (over symmetrical links) as long as QoS is enabled on the uplink, provided that the lower of the two speeds (the uplink) is adequate for transporting both voice and data.

Some service providers for business-class services offer symmetrical links in an effort to compete with Frame Relay providers; 384 kbps/384 kbps and 768 kbps/768 kbps are examples. With no QoS enabled on the service provider edge, this offering is not optimal for the enterprise. An asymmetrical link such as 384 kbps/1.5 kbps is a better choice for the V3PN networks.

Broadband Serialization Mitigation Through TCP Maximum Segment Size Tuning

The majority of broadband deployments are DSL with PPPoE and cable with DOCSIS 1.0. As previously noted, neither of these technologies includes any mechanisms to fragment data packets and interleave voice packets at Layer 2 to minimize the impact of serialization and blocking delay on voice packets (which is a recommended requirement on link speeds £ 768 kbps).

The DOCSIS 1.1 specification for cable includes fragmentation and interleaving support, and DSL providers can implement MLP over ATM (which includes MLP LFI support). However, these do not represent the majority of the currently deployed base of broadband networks.

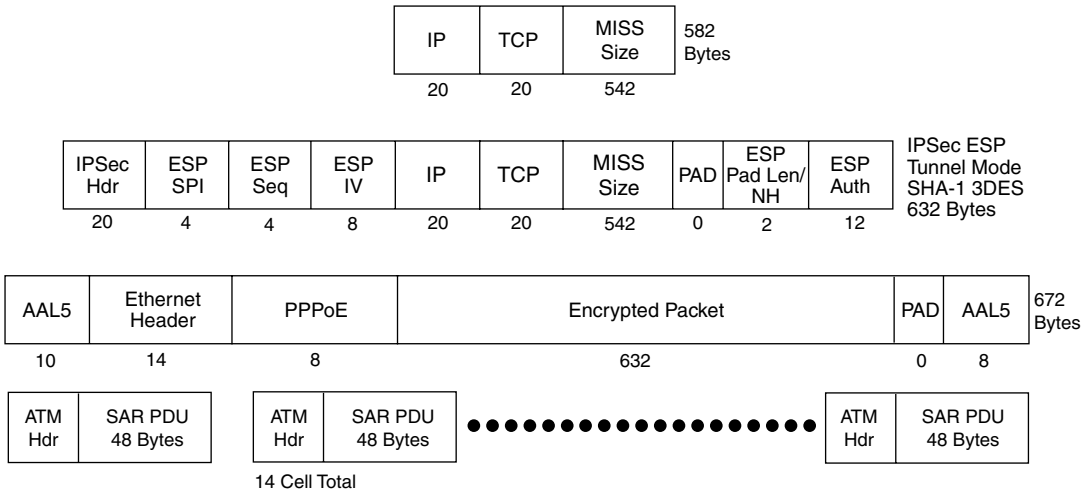
An alternative way to mitigate serialization delay on broadband circuits is provided by adjusting the TCP Maximum Segment Size (MSS). The TCP MSS value influences the resulting size of TCP packets and can be adjusted with the **ip tcp adjust-mss** interface command. Because the majority of large data packets on a network are TCP (normal UDP application packets, such as DNS and NTP, average less than 300 bytes and do not create a significant serialization delay issue), this command effectively can reduce serialization delay in most teleworker scenarios when no Layer 2 fragmentation and interleaving mechanism is available.

Note It is not recommended to run UDP-based video applications on broadband links £ 768 kbps that do not support Layer 2 LFI in a V3PN teleworker scenario. This is because such applications regularly generate large UDP packets that, obviously, are not subject to TCP MSS and thus can cause significant and unmitigatable serialization delays to VoIP packets.

The recommended TCP MSS value of 542 was calculated to eliminate the IPSec crypto algorithm (3DES) padding and the ATM AAL5 padding in DSL implementations (cable implementations have 3DES padding but no AAL5). Therefore, a TCP MSS value of 542 bytes is valid for cable but is optimized for DSL. Figure 6-26 illustrates a TCP packet with an MSS size of 542.

Note Both the ESP and AAL5 pad lengths are 0. A TCP packet with a TCP MSS value of 542 fits exactly into 14 ATM cells, with no wasted bytes because of cell padding.

Figure 6-26 Optimized TCP MSS Value for DSL (542 Bytes)



The evident negative aspect of using this technique to minimize the impact of serialization delay is the decreased efficiency of large data transfers, coupled with an increase in the number of packets per second that the router must switch. The higher packets-per-second rate is not as much of an issue at the remote router as at the headend, where hundreds or thousands of remote connections are concentrated. However, adjusting the TCP MSS value provides a means for the network manager to deploy voice over broadband connections, which do not support any LFI mechanism at rates less than 768 kbps.

Split Tunneling

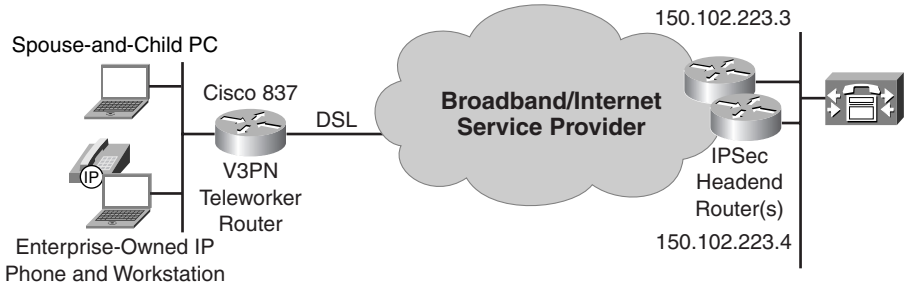
The teleworker is usually not the only user of the residential broadband connection. The worker's spouse and children also might utilize this link from other devices, such as additional PCs or laptops or even gaming units. In such cases, only "work-related" traffic would require encryption and could be given a preferential treatment, through QoS, over general Internet traffic.

In this situation, prioritization could be made based on the destination address of the traffic. One method to accomplish this would be to create a separate bandwidth class (for example, named CORPORATE) and provision this class with a dedicated CBWFQ queue.

Given the sample topology illustrated in Figure 6-27, a suitable QoS configuration is broken down as follows:

- **VOICE**—A single call's VoIP packets assigned to an LLQ, with the bandwidth allocation based on codec and broadband media (DSL or cable).
- **CALL-SIGNALING**—Call-Signaling traffic in a CBWFQ queue, allocated 5 percent.
- **INTERNETWORK**—Internetwork-Control traffic, such as routing protocol updates and IKE keepalives, in a CBWFQ queue, allocated 5 percent.
- **CORPORATE**—All other traffic that is destined to the enterprise through the IPSec tunnel, in a CBWFQ queue, allocated 25 percent.
- **INTERNET/Class-Default**—Traffic to the Internet (outside the IPSec tunnel) defaults to this fair-queued class.

Figure 6-27 Split Tunnel Example Topology



The following configuration fragment provides a means to implement this policy. It assumes that the headend IPSec crypto peers reside on network 150.102.223.0/29. In this example environment, two peers are configured at 150.102.223.3 and 150.102.223.4. Packets within the IPSec tunnel—those matching the CORP-INTRANET access control list (identifying RFC 1918 private addresses that are assumed in this example to represent the intranets behind the headends)—will be encapsulated in an ESP (IPSec-IP protocol 50) IP header. A class-map CORPORATE is configured that references the CORPORATE extended access list, pointing to the VPN headend subnet.

A policy map named V3PN-SPLIT-TUNNEL is created that includes classes for VOICE, INTERNETWORK-CONTROL, and CALL-SIGNALING, along with a bandwidth class for CORPORATE.

In this example, the VOICE class is provisioned for a G.711 codec over a (384-kbps) DSL uplink, allocating 150 kbps (or 40 percent, whichever syntax is preferred) for VoIP. Example 6-8 shows the relevant configuration fragment.

Example 6-8 V3PN-SPLIT-TUNNEL Policy Example

```
!
crypto map SPLIT-TUNNEL-CRYPTO-MAP 1 ipsec-isakmp
 set peer 150.102.223.3
 set peer 150.102.223.4
 set transform-set TS
 match address CORP-INTRANET          ! References CORP-INTRANET ACL
 qos pre-classify                     ! Enables QoS Pre-Classify
!
!
class-map match-all VOICE
  match ip dscp ef                    ! VoIP
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                   ! IP Routing
  match access-group name IKE         ! References ISAKMP ACL
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                   ! New Call-Signaling
  match ip dscp af31                  ! Old Call-Signaling
 class-map match-all CORPORATE
  match access-group name CORPORATE   ! References CORPORATE ACL
!
policy-map V3PN-SPLIT-TUNNEL
  class VOICE
    priority 150                      ! Encrypted G.711 over DSL (PPPoE/AAL5)
  class INTERNETWORK-CONTROL
    bandwidth percent 5               ! Control Plane provisioning
  class CALL-SIGNALING
    bandwidth percent 5               ! Call-Signaling provisioning
  class CORPORATE
    bandwidth percent 25              ! "Work-related" traffic provisioning
```

```
    queue-limit 15                    ! Optional: Anti-Replay Tuning
  class class-default
    fair-queue
    queue-limit 15                    ! Optional: Anti-Replay Tuning
!
...
!
ip access-list extended CORP-INTRANET       ! CORP-INTRANET ACL (RFC 1918)
 permit ip any 10.0.0.0 0.255.255.255
 permit ip any 172.16.0.0 0.15.255.255
 permit ip any 192.168.0.0 0.0.255.255
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp      ! ISAKMP ACL
!
ip access-list extended CORPORATE           ! CORPORATE ACL (VPN Head-ends)
 permit esp any 150.102.223.0 0.0.0.7
!
```

Teleworker V3PN QoS Designs

To review, three deployment models exist for teleworker V3PN scenarios:

- Integrated Unit Model
- Dual-Unit Model
- Integrated Unit + Access Model

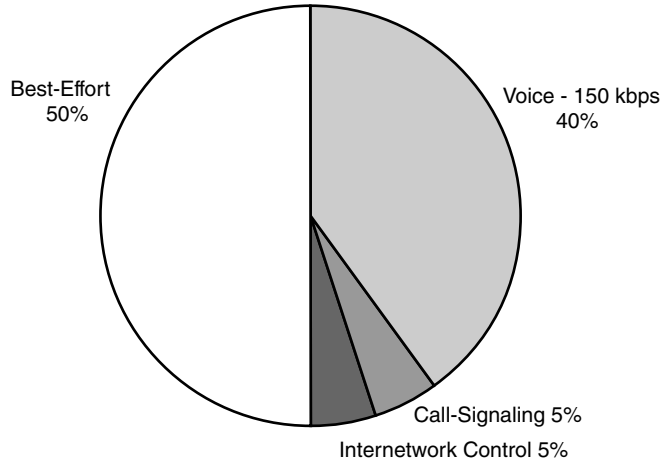
Furthermore, there are two main broadband deployment media: DSL and cable. DSL supports all three teleworker deployment models, but cable supports (at the time of writing) only the Integrated Unit + Access Model.

Integrated Unit/Dual-Unit Models—DSL Design

The key point to remember when provisioning QoS for V3PN over DSL (PPPoE/ATM AAL5) is the significant bandwidth overhead required by these protocols (85 kbps is needed per G.729 call, and 150 kbps is needed per G.711 call).

Some additional points to keep in mind are that since the service policy is being applied to a low-speed ATM PVC, the Tx-ring should be tuned to 3 (as discussed in Chapter 3, "WAN Aggregator QoS Design") and also that because no LFI mechanism exists for DSL, TCP-MSS tuning can be done on the dialer interface to mitigate serialization delay. An Integrated Unit/Dual-Unit V3PN teleworker example for a 384-kbps DSL circuit is illustrated in Figure 6-28 and detailed in Example 6-9.

Figure 6-28 Integrated Unit/Dual-Unit V3PN Teleworker Model for 384-kbps DSL Uplink



Example 6-9 Integrated Unit/Dual-Unit V3PN Teleworker QoS Design Example for a 384-kbps (PPPoE/ATM) DSL Uplink

```
!
 class-map match-all VOICE
  match ip dscp ef                     ! VoIP
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                    ! IP Routing
  match access-group name IKE          ! References ISAKMP ACL
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                    ! New Call-Signaling
  match ip dscp af31                   ! Old Call-Signaling
!
!
 policy-map V3PN-TELEWORKER
  class VOICE
    priority 150                      ! Encrypted G.711 over DSL (PPPoE/AAL5)
  class INTERNETWORK-CONTROL
    bandwidth percent 5               ! Control Plane provisioning
  class CALL-SIGNALING
    bandwidth percent 5               ! Call-Signaling provisioning
  class class-default
    fair-queue
    queue-limit 30                    ! Optional: Anti-Replay Tuning
!
...
!
interface ATM0
 no ip address
 no atm ilmi-keepalive
 dsl operating-mode auto
 dsl power-cutback 0
!
interface ATM0.35 point-to-point
 description Outside PPPoE/ATM DSL Link
 bandwidth 384
 pvc dsl 0/35
  vbr-nrt 384 384
  tx-ring-limit 3                              ! Tx-Ring tuned to 3
  pppoe max-sessions 5
  service-policy output V3PN-TELEWORKER       ! MQC policy applied to PVC
```

```
  pppoe-client dial-pool-number 1
!
...
!
interface Dialer1
 description Dialer for PPPoE
 ip address negotiated
 ip mtu 1492
 encapsulation ppp
 ip tcp adjust-mss 542                        ! TCP MSS value tuned for slow-link
!

...
!
ip access-list extended IKE
 permit udp any eq isakmp any eq isakmp       ! ISAKMP ACL
!
```

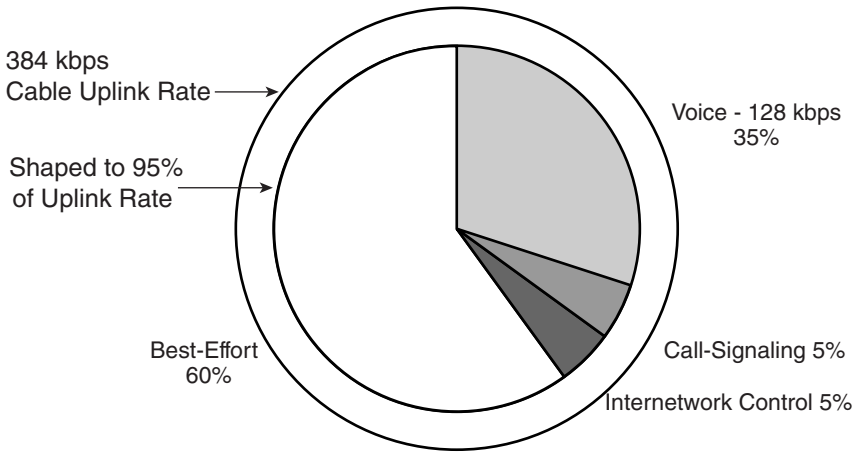
Verification commands:

- **show policy**
- **show policy interface**
- **show atm pvc**

Integrated Unit + Access Model—DSL/Cable Designs

Hierarchical MQC policies are required for Integrated Unit + Access Models to shape and queue (within the shaped rate) on the outbound Ethernet interface. For cable, the shaped rate should be 95 percent of the broadband link's speed (as detailed in Chapter 3, "WAN Aggregator QoS Design" in the "Frame Relay" section ["Committed Information Rate" subsection]). For DSL, the shaped rate should be 70 percent of the uplink's speed (to account for the increased bandwidth overhead required by DSL). Furthermore, on slow-speed (£ 768 kbps) links, TCP-MSS should be tuned on both the inbound and outbound Ethernet interfaces. An Integrated Unit + Access Model for a 384-kbps cable teleworker uplink is shown in Figure 6-29 and detailed in Example 6-10.

Figure 6-29 Integrated Unit + Access V3PN Teleworker Model for 384-kbps Cable Uplink



Example 6-10 Integrated Unit + Access Model—Cable Design Example

```
!
 class-map match-all VOICE
  match ip dscp ef                      ! VoIP
 class-map match-any INTERNETWORK-CONTROL
  match ip dscp cs6                     ! IP Routing
  match access-group name IKE           ! References ISAKMP ACL
 class-map match-any CALL-SIGNALING
  match ip dscp cs3                     ! Old Call-Signaling
  match ip dscp af31                    ! New Call-Signaling
!
!
 policy-map V3PN-TELEWORKER
  class VOICE
    priority 128                        ! Encrypted G.711 over Cable
  class INTERNETWORK-CONTROL
    bandwidth percent 5                 ! Control Plane provisioning
  class CALL-SIGNALING
    bandwidth percent 5                 ! Call-Signaling provisioning
  class class-default
    fair-queue
    queue-limit 30                      ! Optional: Anti-Replay Tuning
!
!
policy-map SHAPE-384-CABLE
 class class-default
   shape average 364800 3640            ! Shapes to 95% of 384 kbps cable link
   service-policy V3PN-TELEWORKER       ! Nested V3PN Teleworker queuing policy
!
...
!
interface Ethernet0
 description Inside Ethernet Interface
 ip tcp adjust-mss 542                          ! TCP MSS value tuned for slow-link
!
interface Ethernet1
 description Outside Ethernet Interface
 ip address dhcp
 ip tcp adjust-mss 542                          ! TCP MSS value tuned for slow-link
 service-policy output SHAPE-384-CABLE          ! Shaper applied to LAN interface
!
```

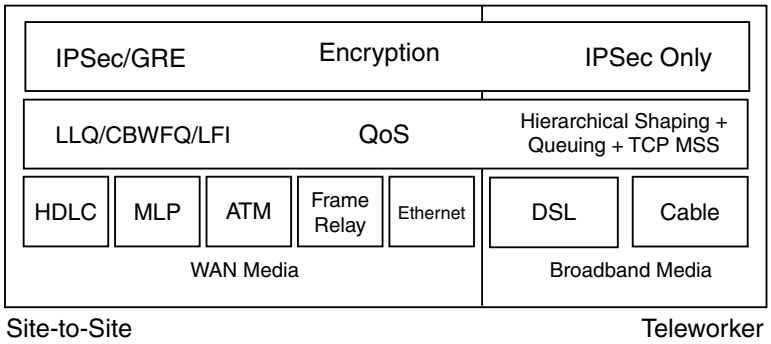
Verification commands:

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

Summary

IPSec VPNs, the most commonly deployed VPN solutions today, are found in three main contexts: site-to-site VPNs, teleworker VPNs, and remote-access VPNs. The overlaying of QoS technologies on top of IPSec VPNs is dubbed V3PN, for voice- and video-enabled Virtual Private Networks. This chapter presented considerations and design recommendations for V3PN deployments in site-to-site and teleworker contexts. A summary of the design recommendations for encryption and QoS for site-to-site and teleworker IPSec V3PNs is illustrated in Figure 6-30.

Figure 6-30 IPSec V3PN Site-to-Site and Teleworker Design Summary Comparison



Some site-to-site considerations that were discussed include the bandwidth implications of various IPSec modes of operations and the incompatibility of cRTP and IPSec. The interrelation of IPSec VPN logical hub-and-spoke topologies and the effect these have on spoke-to-spoke delay budgets also were examined. Subsequently, the ToS byte preservation mechanism was overviewed along with the QoS Pre-Classify Cisco IOS Software feature, which allows for the classification of packets already encrypted (on the same router) through ACLs. QoS and Anti-Replay implications then were discussed, illustrating how QoS policies that reorder packets potentially can exacerbate Anti-Replay drops (an IPSec message-integrity mechanism). Techniques for minimizing such undesired QoS/Anti-Reply interaction effects, such as reducing the queue length of data queues, were presented. Next, the need for control plane provisioning was highlighted, along with basic designs for doing so.

Several site-to-site QoS models were detailed, ranging from a six-class V3PN QoS model to a complex 11-class V3PN QoS Baseline model. WAN aggregator considerations specific to IPSec VPN deployments were examined next, including QoS provisioning for IPSec over private WANs, per-tunnel hierarchical shaping and queuing, and recommendations for decoupled VPN headend/WAN aggregation deployment models, where encryption and QoS are performed on different routers.

Attention then shifted to teleworker scenarios and the three main teleworker deployment models: the Integrated Unit Model, the Dual-Unit Model, and the Integrated Unit + Access Model. The two main broadband media types, DSL and cable, were broken down to ascertain the bandwidth-provisioning implications of each media. Neither DSL nor (DOCSIS 1.0) cable includes any mechanism for serialization delay mitigation, so TCP maximum segment-size tuning was considered as an alternative mechanism to achieve this. Split-tunneling designs, to address spouse-and-child requirements, were introduced.

Teleworker V3PN designs then were detailed for Integrated Unit and Dual-Unit models over DSL, in addition to an Integrated Unit + Access Model solution for cable.

References

Standards

- RFC 1321, “The MD5 Message-Digest Algorithm”: http://www.ietf.org/rfc/rfc1321.txt.
- RFC 1918, “Address Allocation for Private Internets”: http://www.ietf.org/rfc/rfc1918.txt.
- RFC 2104, “HMAC: Keyed-Hashing for Message Authentication”: http://www.ietf.org/rfc/rfc2104.txt.

- RFC 2401, “Security Architecture for the Internet Protocol”: http://www.ietf.org/rfc/rfc2401.txt.
- RFC 2402, “IP Authentication Header”: http://www.ietf.org/rfc/rfc2402.txt.
- RFC 2403, “The Use of HMAC-MD5-96 Within ESP and AH”: http://www.ietf.org/rfc/rfc2403.txt.
- RFC 2404, “The Use of HMAC-SHA-1-96 within ESP and AH”: http://www.ietf.org/rfc/rfc2404.txt.
- RFC 2405, “The ESP DES-CBC Cipher Algorithm with Explicit IV”: http://www.ietf.org/rfc/rfc2405.txt.
- RFC 2406, “IP Encapsulating Security Payload (ESP)”: http://www.ietf.org/rfc/rfc2406.txt.
- RFC 2407, “The Internet IP Security Domain of Interpretation for ISAKMP”: http://www.ietf.org/rfc/rfc2407.txt.
- RFC 2408, “Internet Security Association and Key Management Protocol (ISAKMP)”: http://www.ietf.org/rfc/rfc2408.txt.
- RFC 2409, “The Internet Key Exchange (IKE)”: http://www.ietf.org/rfc/rfc2409.txt.
- RFC 2410, “The NULL Encryption Algorithm and Its Use with IPSec”: http://www.ietf.org/rfc/rfc2410.txt.
- RFC 2411, “IP Security Document Roadmap”: http://www.ietf.org/rfc/rfc2411.txt.
- RFC 2412, “The OAKLEY Key Determination Protocol”: http://www.ietf.org/rfc/rfc2412.txt.

Books

- Kaeo, Merike. *Designing Network Security.* Indianapolis: Cisco Press, 2003.
- Malik, Saadat. *Network Security Principles and Practices*. Indianapolis: Cisco Press, 2002.
- Mason, Andrew. *Cisco Secure Virtual Private Networks.* Indianapolis: Cisco Press, 2001.

Cisco IOS Documentation

- IP Security and Encryption overview (Cisco IOS Release 12.2): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fsecur_c/fipsenc/scfencov.htm.
- Configuring IPSec network security (Cisco IOS Release 12.2): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fsecur_c/fipsenc/scfipsec.htm.
- Configuring Internet Key Exchange Security Protocol (Cisco IOS Release 12.2): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fsecur_c/fipsenc/scfike.htm.
- Prefragmentation for IPSec VPNs (Cisco IOS Release 12.2[13]T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftprefrg.htm.
- Quality of service for Virtual Private Networks (Cisco IOS Release 12.2[2]T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t2/ftqosvpn.htm.

- Low-latency queuing (LLQ) for IPSec encryption engines (Cisco IOS Release 12.2[13]T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/llqfm.htm.
- IPSec NAT transparency (Cisco IOS Release 12.2[13]T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/ftipsnat.htm.
- IPSec and quality of service feature (Cisco IOS Release 12.3[8]T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123newft/123t/123t_8/gtqosips.htm.
- Configuring broadband access (Cisco IOS Release 12.2): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fwan_c/wcfppp.htm.
- PPP over Ethernet Client (Cisco IOS Release 12.2[2]T): http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t2/ftpppoec.htm.