QOS

Implementing Cisco Quality of Service

Volume 2

Version 2.2

Student Guide

Editorial, Production, and Graphic Services: 06.28.06

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

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

Asia Pacific Headquarters
Cisco Systems, Inc.
168 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

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

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

© 2006 Cisco Systems, Inc. All rights reserved. CCSP, CCVP, the Cisco Square Bridge logo, Follow Me Browsing, and StackWise are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn, and iQuick Study are service marks of Cisco Systems, Inc.; and Access Registrar, Aironet, BPX, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, FormShare, GigaDrive, GigaStack, HomeLink, Internet Quotient, IOS, IP/TV, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, LightStream, Linksys, MeetingPlace, MGX, the Networkers logo, Networking Academy, Network Registrar, *Packet,* PIX, Post-Routing, Pre-Routing, ProConnect, RateMUX, ScriptShare, SlideCast, SMARTnet, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0601R)

DISCLAIMER WARRANTY: THIS CONTENT IS BEING PROVIDED "AS IS." CISCO MAKES AND YOU RECEIVE NO WARRANTIES IN CONNECTION WITH THE CONTENT PROVIDED HEREUNDER, EXPRESS, IMPLIED, STATUTORY OR IN ANY OTHER PROVISION OF THIS CONTENT OR COMMUNICATION BETWEEN CISCO AND YOU. CISCO SPECIFICALLY DISCLAIMS ALL IMPLIED WARRANTIES, INCLUDING WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT AND FITNESS FOR A PARTICULAR PURPOSE, OR ARISING FROM A COURSE OF DEALING, USAGE OR TRADE PRACTICE. This learning product may contain early release content, and while Cisco believes it to be accurate, it falls subject to the disclaimer above.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Table of Contents

Volume 2

Congestion Management ***5-1***

Overview 5-1
Module Objectives 5-1

Introducing Queuing ***5-3***

Overview 5-3
Objectives 5-3
Congestion and Queuing 5-4
Queuing Algorithms 5-7
First In First Out 5-8
Priority Queuing 5-9
Round Robin 5-10
Weighted Round Robin 5-11
Deficit Round Robin 5-13
Summary 5-14

Understanding Queuing Implementations ***5-15***

Overview 5-15
Objectives 5-15
Queuing Components 5-16
Hardware Queue (TxQ) Size 5-19
Congestion on Software Interfaces 5-21
Queuing Implementations in Cisco IOS 5-22
Summary 5-23

Configuring FIFO and WFQ ***5-25***

Overview 5-25
Objectives 5-25
FIFO Queuing 5-26
Weighted Fair Queuing 5-28
WFQ Classification 5-31
WFQ Insertion and Drop Policy 5-34
WFQ Scheduling 5-36
Benefits and Drawbacks of WFQ 5-44
Configuring WFQ 5-45
Monitoring WFQ 5-48
Summary 5-50

Configuring CBWFQ and LLQ ***5-51***

Overview 5-51
Objectives 5-51
CBWFQ and LLQ 5-52
Class-Based Weighted Fair Queuing 5-53
CBWFQ Architecture 5-54
CBWFQ Benefits 5-60
Configuring and Monitoring CBWFQ 5-61
Example: Configuration of FIFO Queuing 5-63
Example: Configuration of WFQ Queuing 5-63
Low-Latency Queuing 5-65
LLQ Architecture 5-67
LLQ Benefits 5-68
Configuring and Monitoring LLQ 5-69
Example: Calculating LLQ Bandwidth Required for VoIP 5-71
Summary 5-74
References 5-75

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring LAN Congestion Management 5-77

Overview 5-77
Objectives 5-77
Queuing on Catalyst Switches 5-78
Weighted Round Robin 5-84
Configuring CoS-to-Queue Mappings for PQ on Catalyst 2950 Switches 5-86
Configuring WRR on Catalyst 2950 Switches 5-87
Monitoring Queuing on Catalyst 2950 Switches 5-89
Summary 5-92
Module Summary 5-93
References 5-94
Module Self-Check 5-95
Module Self-Check Answer Key 5-97

Congestion Avoidance 6-1

Overview 6-1
Module Objectives 6-1

Introducing Congestion Avoidance 6-3

Overview 6-3
Objectives 6-3
Behavior of TCP Senders and Receivers 6-4
Example: Windowing in TCP 6-5
Congestion and TCP 6-6
Managing Interface Congestion with Tail Drop 6-9
Tail Drop Limitations 6-10
Summary 6-13

Introducing RED 6-15

Overview 6-15
Objectives 6-15
Random Early Detection 6-16
RED Profiles 6-17
RED Modes 6-19
TCP Traffic Before and After RED 6-20
Applying Congestion Avoidance 6-22
Summary 6-23

Configuring Class-Based Weighted RED 6-25

Overview 6-25
Objectives 6-25
Weighted Random Early Detection 6-26
WRED Profiles 6-30
Configuring CB-WRED 6-35
Example: CBWFQ Using IP Precedence with CB-WRED 6-40
Configuring DSCP-Based CB-WRED 6-42
Example: CB-WRED Using DSCP with CBWFQ 6-44
Monitoring CB-WRED 6-46
Summary 6-48

Configuring Explicit Congestion Notification 6-49

Overview 6-49
Objectives 6-49
Explicit Congestion Notification 6-50
ECN Field Defined 6-52
ECN and WRED 6-53
Configuring ECN-Enabled WRED 6-55
Monitoring ECN-Enabled WRED 6-56

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary 6-59
Module Summary 6-60
References 6-61
Module Self-Check 6-62
Module Self-Check Answer Key 6-66

Traffic Policing and Shaping 7-1

Overview 7-1
Module Objectives 7-1

Understanding Traffic Policing and Shaping 7-3

Overview 7-3
Objectives 7-4
Traffic Policing and Shaping Overview 7-5
Why Use Traffic Conditioners? 7-6
Example: Traffic Policing 7-8
Example: Traffic Shaping 7-9
Policing vs. Shaping 7-10
Measuring Traffic Rates 7-11
Example: Token Bucket as a Piggy Bank 7-12
Single Token Bucket Class-Based Policing 7-14
Dual Token Bucket Class-Based Policing 7-15
Dual-Rate Token Bucket Class-Based Policing 7-17
Example: Dual-Rate Token Bucket as a Piggy Bank 7-19
Class-Based Traffic Shaping 7-20
Cisco IOS Traffic Policing and Shaping Mechanisms 7-21
Applying Traffic Conditioners 7-23
Summary 7-24

Configuring Class-Based Policing 7-25

Overview 7-25
Objectives 7-25
Class-Based Policing Overview 7-26
Configuring Single-Rate Class-Based Policing 7-28
Example: Single Rate, Dual Token Bucket Class-Based Policing 7-30
Example: Multiaction Class-Based Policing 7-31
Configuring Dual-Rate Class-Based Policing 7-32
Example: Dual-Rate Class-Based Policing 7-33
Configuring Percentage-Based Class-Based Policing 7-34
Example: Configuring Percentage-Based Class-Based Policing 7-35
Monitoring Class-Based Policing 7-36
Summary 7-37

Configuring Class-Based Shaping 7-39

Overview 7-39
Objectives 7-39
Class-Based Shaping Overview 7-40
Traffic Shaping Methods 7-41
Configuring Class-Based Shaping 7-42
Example: Average Rate, Peak Rate 7-43
Example: Class-Based Shaping with CBWFQ 7-45
Example: Class-Based Shaping Hierarchical Policy Maps 7-46
Monitoring Class-Based Shaping 7-47
Summary 7-49

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Class-Based Shaping on Frame Relay Interfaces ***7-51***

Overview 7-51
Objectives 7-51
Frame Relay Refresher 7-52
Frame Relay Congestion Control 7-53
Frame Relay Congestion Adaptation 7-54
FECN-to-BECN Propagation 7-55
Configuring Frame Relay Adaptive Class-Based Shaping 7-56
Example: Class-Based Shaping with Frame Relay Adaptation 7-57
Monitoring MQC-Based Frame Relay Traffic Shaping 7-58
MQC-Based Frame Relay Traffic Shaping Example on Multipoint Main Interface 7-60
Summary 7-63

Frame Relay Voice-Adaptive Traffic Shaping and Fragmentation ***7-65***

Overview 7-65
Objectives 7-66
Frame Relay Voice-Adaptive Traffic Shaping and Fragmentation 7-67
Benefits in Deploying FR-VATS 7-68
Prerequisites for Deploying FR-VATS 7-69
Supported Platforms 7-70
Frame Relay Voice-Adaptive Traffic Shaping and Fragmentation Operation 7-71
Configuring the Frame Relay Voice-Adaptive Feature 7-72
Monitoring the Frame Relay Voice-Adaptive Feature 7-76
Summary 7-78
Module Summary 7-79
References 7-80
Module Self-Check 7-81
Module Self-Check Answer Key 7-83

Link Efficiency Mechanisms ***8-1***

Overview 8-1
Module Objectives 8-1

Understanding Link Efficiency Mechanisms ***8-3***

Overview 8-3
Objectives 8-3
Link Efficiency Mechanisms Overview 8-4
Example: Indexing Operations 8-6
L2 Payload Compression 8-7
Example: L2 Payload Compression Results 8-8
Header Compression 8-9
Example: Header Compression Results 8-10
Large Packets “Freeze Out” Voice on Slow WAN Links 8-11
Link Fragmentation and Interleaving 8-13
Applying Link Efficiency Mechanisms 8-14
Summary 8-15

Configuring Class-Based Header Compression ***8-17***

Overview 8-17
Objectives 8-17
Header Compression Overview 8-18
Example: RTP Header Compression 8-19
Class-Based TCP Header Compression 8-20
Example: Class-Based TCP Header Compression 8-21
Class-Based RTP Header Compression 8-23
Example: Class-Based RTP Header Compression 8-24
Configuring Class-Based Header Compression 8-26
Example: Configuring Class-Based TCP Header Compression 8-27
Example: Configuring Class-Based RTP Header Compression 8-28

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring Class-Based Header Compression 8-29
Summary 8-30

Configuring Link Fragmentation and Interleaving 8-31

Overview 8-31
Objectives 8-31
Fragmentation Options 8-32
Serialization Delay and Fragment Sizing 8-33
Example: Determining the Proper Fragment Size 8-34
Configuring MLP with Interleaving 8-35
Example: MLP with Interleaving 8-37
Monitoring MLP with Interleaving 8-38
FRF.12 Frame Relay Fragmentation 8-40
Configuring FRF.12 Frame Relay Fragmentation 8-42
Example: FRF.12 Frame Relay Fragmentation 8-43
Monitoring FRF.12 Frame Relay Fragmentation 8-45
Summary 8-47
Module Summary 8-48
References 8-49
Module Self-Check 8-50
Module Self-Check Answer Key 8-53

QoS Best Practices 9-1

Overview 9-1
Module Objectives 9-1

Understanding Traffic Classification Best Practices 9-3

Overview 9-3
Objectives 9-3
QoS Best Practices 9-4
Example: Cisco IOS QoS Tools Summary 9-8
Voice, Video, and Data QoS Requirements 9-9
Example: G.711 Voice Bearer Bandwidth Requirement Calculation 9-11
Example: Calculating the Bandwidth Requirement for a 384-kbps Videoconference Stream 9-13
Example: RCL Enterprise 9-15
QoS Requirements Summary 9-16
Example: QoS Requirements of the Major Applications Category 9-16
Traffic Classification 9-17
Example: LLQ Bandwidth Allocation 9-19
Example: LLQ Example on the Enterprise WAN Edge Router 9-22
Enterprise to Service Provider QoS Class Mapping 9-23
Example: Remapping Enterprise-Managed CE Traffic Classes into Traffic Classes Offered by Service Provider 9-24
Summary 9-26

Deploying End-to-End QoS 9-27

Overview 9-27
Objective 9-27
QoS Service Level Agreements 9-28
Deploying End-to-End QoS 9-34
Enterprise Campus QoS Implementations 9-36
WAN Edge (CE/PE) QoS Implementations 9-43
Service Provider Backbone QoS Implementations 9-54
Summary 9-62

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Providing QoS for Security *9-63*

Overview 9-63
Objective 9-63
Business Security Threat 9-64
Increasing Occurrence of Network Attacks 9-65
The Impact of an Internet Worm Attack 9-66
QoS Tools and Tactics for Security 9-67
Control Plane Policing 9-69
Data Plane Policing 9-70
NBAR Known-Worm Policing 9-74
Integrating Security Through QoS 9-75
Summary 9-76
Module Summary 9-77
References 9-78
Module Self-Check 9-79
Module Self-Check Answer Key 9-82

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module 5

Congestion Management

Overview

Congestion can occur in many different locations within a network and is the result of many factors, including oversubscription, insufficient packet buffers, traffic aggregation points, network transit points, and speed mismatches (such as at LAN-to-WAN links). Simply increasing link bandwidth is not adequate to solve the congestion issue, in most cases. Aggressive traffic can fill interface queues and starve more fragile flows such as voice and interactive traffic. The results can be devastating for delay-sensitive traffic types, making it difficult to meet the service-level requirements these applications require. Fortunately, there are many congestion management techniques available on Cisco IOS platforms, which provide you with an effective means to manage software queues and to allocate the required bandwidth to specific applications when congestion exists. This module examines the components of queuing systems and the different congestion management mechanisms available on Cisco IOS devices.

Module Objectives

Upon completing this module, you will be able to use Cisco quality of service (QoS) queuing mechanisms to manage network congestion. This ability includes being able to meet these objectives:

- Explain the operation of basic queuing algorithms, including FIFO, priority, and round-robin queuing
- Describe hardware and software queuing on a network device
- Configure WFQ to manage congestion
- Configure CBWFQ and LLQ to manage congestion
- Configure WRR on a Catalyst switch to manage LAN congestion

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 1

Introducing Queuing

Overview

Queuing algorithms are one of the primary ways to manage congestion in a network. Network devices handle an overflow of arriving traffic by using a queuing algorithm to sort traffic and determine a method of prioritizing the traffic onto an output link. Each queuing algorithm was designed to solve a specific network traffic problem and has a particular effect on network performance. This lesson describes the basic queuing algorithms.

Objectives

Upon completing this lesson, you will be able to explain the operation of basic queuing algorithms, including FIFO, priority, and round-robin queuing. This ability includes being able to meet these objectives:

- Explain the need for congestion management mechanisms
- List the different queuing algorithms
- Describe the FIFO queuing algorithm
- Describe the priority queuing algorithm
- Describe the round-robin queuing algorithm
- Describe the WRR queuing algorithm
- Describe the deficit round-robin queuing algorithm

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Congestion and Queuing

This topic describes the need for congestion management mechanisms.

Congestion can occur anywhere within a network where speed mismatches (for example, a GigabitEthernet link feeding a FastEthernet link), aggregation (for example, multiple GigabitEthernet links feeding an upstream GigabitEthernet), or confluence (the flowing together of two or more traffic streams).

Queuing algorithms are used to manage congestion. Many algorithms have been designed to serve different needs. A well-designed queuing algorithm will provide some bandwidth and delay guarantees to priority traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Congestion and Queuing: Speed Mismatch

- **Speed mismatches are the most typical cause of congestion.**
- **Possibly persistent when going from LAN to WAN.**
- **Usually transient when going from LAN to LAN.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-3

Speed mismatches are the most typical cause of congestion in a network.

Speed mismatches are most common when traffic moves from a high-speed LAN environment (1000 Mbps or higher) to lower-speed WAN links (1 or 2 Mbps). Speed mismatches are also common in LAN-to-LAN environments when, for example, a 1000-Mbps link feeds into a 100-Mbps link.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The most common site of congestion is at points of aggregation in a network.

Typical points of aggregation occur in WANs when multiple remote sites feed back into a central services site.

In a LAN environment, congestion resulting from aggregation often occurs at the distribution layer of networks, where the different access layer devices feed traffic to the distribution-level switches.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Queuing Algorithms

This topic describes the different queuing algorithms.

Queuing Algorithms

- FIFO
- PQ
- Round robin
- WRR
- DRR

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-5

Key queuing algorithms include the following:

- **FIFO:** The simplest algorithm
- **Priority queuing (PQ):** Allows certain traffic to be strictly-prioritized
- **Round robin:** Allows several traffic flows to share bandwidth
- **Weighted round robin (WRR):** Allows sharing of bandwidth with preferential treatment
- **Deficit round robin (DRR):** Resolves problem with some WRR implementations

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

First In First Out

This topic describes the FIFO queuing algorithm.

FIFO is the simplest queuing algorithm. Packets are placed into a single queue and serviced in the order they were received.

All individual queues are, in fact, FIFO queues. Other queuing methods rely upon FIFO as the congestion management mechanism for single queues, while using multiple queues to perform more advanced functions such as prioritization.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Priority Queuing

This topic describes the priority queuing algorithm.

The PQ algorithm is also quite simple. Each packet is assigned a priority and placed into a hierarchy of queues based on priority. When there are no more packets in the highest queue, the next lower queue is serviced. Packets are then dispatched from the next highest queue until either the queue is empty or another packet arrives for a higher PQ.

Packets will be dispatched from a lower queue only when all higher priority queues are empty. If a packet arrives for a higher queue, the packet from the higher queue is dispatched before any packets in lower-level queues.

The problem with PQ is that queues with lower priority can "starve" if a steady stream of packets continues to arrive for a queue with a higher priority. Packets waiting in the lower priority queues may never be dispatched.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Round Robin

This topic describes the round-robin queuing algorithm.

With round-robin queuing, one packet is taken from each queue and then the process repeats.

If all packets are the same size, all queues share the bandwidth equally. If packets being put into one queue are larger, that queue will receive a larger share of bandwidth.

No queue will “starve” with round robin because all queues receive an opportunity to dispatch a packet every round.

A limitation of round robin is the inability to prioritize traffic.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Weighted Round Robin

This topic describes the WRR queuing algorithm.

The WRR algorithm was developed to provide prioritization capabilities for round robin.

In WRR, packets are assigned a class (voice, file transfer, and so on) and placed into the queue for that class of service. Packets are accessed round-robin style, but queues can be given priorities called "weights." For example, in a single round, four packets from a high-priority class might be dispatched, followed by two from a middle-priority class, and then one from a low-priority class.

Some implementations of the WRR algorithm will dispatch a configurable number of bytes during each round.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Weighted Round Robin (Cont.)

Problem with WRR

- **Some implementations of WRR dispatch a configurable number of bytes (threshold) from each queue for each round—several packets can be sent in each turn.**
- **The router is allowed to send the entire packet even if the sum of all bytes is more than the threshold.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-10

Some implementations of the weighted round robin algorithm provide preferential service by dispatching a configurable number of bytes each round rather than a number of packets. The Cisco custom queuing (CQ) mechanism is an example of this implementation.

The figure illustrates a unique-case scenario of the WRR algorithm, in which these parameters were used to implement WRR queuing on an interface:

- Maximum Transmission Unit (MTU) of the interface is 1500 bytes.
- The byte count to be sent for the queue in each round is 3000 bytes (twice the MTU).

The example shows how the router first sent two packets with a total size of 2999 bytes. Because this is still within the limit (3000), the router can send the next packet (MTU-sized). The result was that the queue received almost 50 percent more bandwidth in this round than it should have received.

This is one of the drawbacks of WRR queuing—it does not allocate bandwidth accurately.

The limit or weight of the queue is configured in bytes. The accuracy of WRR queuing depends on the weight (byte count) and the MTU. If the ratio between the byte count and the MTU is too small, WRR queuing will not allocate bandwidth accurately. If the ratio between the byte count and the MTU is too large, WRR queuing will cause long delays.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Deficit Round Robin

This topic describes the deficit round robin queuing algorithm.

Deficit Round Robin

- **Solves problem with some implementations of WRR**
- **Keeps track of the number of "extra" bytes dispatched in each round—the "deficit"**
- **Adds the "deficit" to the number of bytes dispatched in the next round**
- **Problem resolved with deficit round robin:**
  - **Threshold of 3000**
  - **Packet sizes of 1500, 1499, and 1500**
  - **Total sent in round = 4499 bytes**
  - **Deficit = (4499 – 3000) = 1499 bytes**
  - **On the next round send only the (threshold – deficit) = (3000 – 1499) = 1501 bytes**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-11

Deficit round robin is an implementation of the WRR algorithm developed to resolve the inaccurate bandwidth allocation problem with WRR. The Cisco modified deficit round robin (MDRR) method used on the Cisco 12000 series is an implementation of deficit round robin.

Deficit round robin uses a deficit counter to track the number of "extra" bytes dispatched over the number of bytes that was to be configured to be dispatched each round. During the next round, the number of "extra" bytes—the deficit—is effectively subtracted from the configurable number of bytes that are dispatched.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Congestion can occur at any point in the network, but particularly at points of speed mismatches and traffic aggregation.**
- **Three basic queuing algorithms are used to manage congestion: FIFO, priority, and round-robin queuing.**
- **FIFO is the simplest queuing algorithm.**
- **Priority queuing allows for the prioritization of traffic through the use of multiple queues, but can starve lower-priority queues.**
- **Round-robin queuing uses multiple queues to provide equal access to all queues.**
- **Weighted round robin offers priority access to multiple queues by assigning "weights" to queues, but some implementations may provide inaccurate access to some queues.**
- **Deficit round-robin queuing solves the inaccuracy problem with round robin by keeping a "deficit" count.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-12

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 2

Understanding Queuing Implementations

Overview

Queuing technologies are one of the primary ways to manage congestion in a network. Network administrators and network designers benefit from an understanding of how the queuing algorithms are implemented. This lesson describes the relationship of queuing components and how this affects the overall queuing performance.

Objectives

Upon completing this lesson, you will be able to describe hardware and software queuing on a network device. This ability includes being able to meet these objectives:

- Describe the primary components of a queuing mechanism
- Explain the significance of the size of the hardware queue
- Explain how congestion occurs on software interfaces
- Describe the basic software queuing technologies used on Cisco network devices

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Queuing Components

This topic describes the primary components of a queuing mechanism.

Queuing on routers is necessary to accommodate bursts when the arrival rate of packets is greater than the departure rate, usually because of one of these two reasons:

- Input interface is faster than the output interface
- Output interface is receiving packets coming in from multiple other interfaces

Initial implementations of queuing used a single FIFO strategy. More complex scheduling mechanisms were introduced when special requirements needed routers to differentiate between packets of different levels of importance.

Queuing was split into these two parts:

- **Hardware queue:** Uses FIFO strategy, which is necessary for the interface drivers to transmit packets one by one. The hardware queue is sometimes referred to as the transmit queue (TxQ).
- **Software queue:** Schedules packets into the hardware queue based on the QoS requirements.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The figure illustrates these actions that have to be taken before a packet can be transmitted:

- Most queuing mechanisms include classification of packets.
- After a packet is classified, a router has to determine whether it can put the packet into the queue or drop the packet. Most queuing mechanisms will drop a packet only if the corresponding queue is full (tail drop). Some mechanisms use a more intelligent dropping scheme, such as weighted fair queuing (WFQ), or a random dropping scheme, such as weighted random early detection (WRED).
- If the packet is allowed to be enqueued, it will be put into the FIFO queue for that particular class.
- Packets are then taken from the individual per-class queues and put into the hardware queue.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The Software Queue

- **Generally, a full hardware queue indicates interface congestion, and software queuing is used to manage it.**
- **When a packet is being forwarded, the router will bypass the software queue if the hardware queue has space in it (no congestion).**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-3

Software queuing is implemented when the interface is congested; the software queuing system is bypassed whenever there is room in the hardware queue.

The software queue is, therefore, only used when data must wait to be placed into the hardware queue.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Hardware Queue (TxQ) Size

This topic describes the significance of the size of the hardware queue.

Hardware Queue (TxQ) Size

- **Routers determine the length of the hardware queue based on the configured bandwidth of the interface.**
- **The length of the hardware queue can be adjusted with the** tx-ring-limit **command.**
- **Reducing the size of the transmit ring has two benefits:**
  - **It reduces the maximum amount of time that packets wait in the FIFO queue before being transmitted.**
  - **It accelerates the use of QoS in the Cisco IOS software.**
- **Improper tuning of the hardware queue may produce undesirable results:**
  - **Long TxQ may result in poor performance of the software queue.**
  - **Short TxQ may result in a large number of interrupts, which cause high CPU utilization and low link utilization.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-4

The double queuing strategy (software and hardware queue) has its impacts on the results of overall queuing. Software queues serve a valuable purpose. If the hardware queue is too long, it will contain a large number of packets scheduled in the FIFO fashion. A long FIFO hardware queue most likely defeats the purpose of the QoS design requiring a certain complex software queuing system (for example, custom queuing [CQ]).

So why use the hardware queue at all? Or why not just set its length to one? That would force all packets to go through the software queue and be scheduled one by one to the interface for transmission. This approach has these drawbacks:

- Each time a packet is transmitted, the interface driver interrupts the CPU and requests more packets to be delivered into its hardware queue. Some queuing mechanisms have complex scheduling that takes time to deliver more packets. The interface does not send anything during that time (link utilization is decreased) if the hardware queue is empty, because its maximum size is one.
- The CPU schedules packets one by one instead of many at the same time (in the same interrupt interval), which increases the CPU utilization.

The default TxQ size is determined by the Cisco IOS software, based on the bandwidth of the media, and should be fine for most queuing implementations. Some platforms and QoS mechanisms will automatically adjust the TxQ size to an appropriate value. Faster interfaces have longer hardware queues because they produce less delay. Slower interfaces have shorter hardware queues to prevent too much delay in the worst-case scenario in which the entire hardware queue is full of MTU-sized packets.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Note Refer to Cisco IOS software configuration documentation for more information.

The transmit ring serves as a staging area for packets in line to be transmitted. The router needs to enqueue a sufficient number of packets on the transmit ring and ensure that the interface driver has packets with which to fill available cell timeslots.

The primary reason to tune the transmit ring is to reduce latency caused by queuing. On any network interface, queuing forces a choice between latency and the amount of burst that the interface can sustain. Larger queue sizes sustain longer bursts while increasing delay. Tune the size of a queue when traffic is experiencing unnecessary delay.

The size of the transmit ring must be small enough to avoid introducing latency because of queuing. The transmit ring must also be large enough to avoid drops and a resulting impact to TCP-based flows. Queuing on the transmit ring introduces a serialization delay that is directly proportional to the depth of the ring. An excessive serialization delay can impact latency budgets for delay-sensitive applications such as voice. Cisco recommends reducing the size of the transmit ring for virtual circuits (VCs) carrying voice. Select a value based on the amount of serialization delay, expressed in seconds, introduced by the transmit ring. Use the following formula:

```
((P*8)*D)/S

P = Packet size in bytes. Multiply by eight to convert to
bits.

D = Transmit-ring depth.

S = Speed of the VC in bps.
```

Note IP packets on the Internet are typically one of three sizes: 64 bytes (for example, control messages), 1500 bytes (for example, file transfers), or 256 bytes (all other traffic). These values produce a typical overall Internet packet size of 250 bytes.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Congestion on Software Interfaces

This topic describes how congestion occurs on software interfaces.

Congestion on Software Interfaces

- **Subinterfaces and software interfaces do not have their own separate Tx ring, therefore no congestion can occur.**
  - **Dialers, tunnels, Frame Relay subinterfaces**
  - **They congest when their main hardware interface Tx ring congests**
- **The Tx-ring state (full, not full) is, therefore, an indication of congestion for software interfaces.**
- **Only hardware interfaces have a Tx ring.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-5

Subinterfaces and software interfaces do not have their own separate transmit (Tx) ring; therefore, no congestion can occur. These interface types include dialers, tunnels, and Frame Relay subinterfaces, and will only congest when their main hardware interface Tx ring congests. The Tx ring state is an indication of congestion for software interfaces.

Software interfaces can implement a software shaping queue when traffic shaping is configured on the Frame Relay subinterface, and, in this case, the main interface software queue will be a FIFO queue.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Queuing Implementations in Cisco IOS

This topic describes the basic software queuing technologies used on Cisco network devices.

Queuing Implementations in Cisco IOS

- **Priority queuing**
  - **Implementation of priority queuing**
  - **Four queues (high, medium, normal, low)**
- **Custom queuing**
  - **Implementation of weighted round robin**
  - **Up to 16 queues**
  - **Threshold based on number of bytes**
  - **Configurable priority queues**
  - **Inaccurate bandwidth allocation due to threshold issue with weighted round robin**
- **Modified deficit round robin**
  - **Deficit round robin with a priority queue for Cisco 12xxx routers**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-6

The figure shows these available software queuing technologies:

- Priority queuing (PQ):
  - A Cisco implementation of the priority queuing algorithm
  - Allows four queues to be used for prioritization (high, medium, normal, low)
  - Allows for a variety of classification, including source IP address, destination IP address, IP precedence, and differentiated services code point (DSCP)
- Custom queuing (CQ):
  - A Cisco implementation of weighted round robin (WRR)
  - Allows up to 16 queues to be used for traffic classification
  - Allows for a variety of classification including: source IP address, destination IP address, IP Precedence, and DSCP
  - Tail drop used within each individual queue
- Modified deficit round robin (MDRR):
  - A Cisco implementation of deficit round robin
  - Available only on the Cisco 12000 series routers

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Each physical interface has a hardware and a software queuing system. The hardware queuing system uses FIFO, while the software queuing system can be configured depending on the platform and IOS version.**
- **The length of the hardware queue has a significant impact on performance and can be configured on a router with the** tx-ring-limit **command.**
- **Software interfaces have no queues; they congest only when their hardware interface congests.**
- **Cisco offers implementations of basic queuing algorithms: priority queuing, custom queuing, and modified deficit round robin.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-7

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 3

Configuring FIFO and WFQ

Overview

FIFO and weighted fair queuing (WFQ) are the two primary default queuing mechanisms that are implemented on Cisco routers. WFQ was developed to resolve some of the problems, such as queue starvation, delay, and jitter, resulting from the use of basic queuing methods. WFQ dynamically divides available bandwidth by a calculation based on the total number of flows and the weight of each given flow. Bandwidth cannot be guaranteed, because the number of flows are constantly changing and therefore so is the allocated bandwidth to each flow. This lesson describes how to configure WFQ to manage congestion.

Objectives

Upon completing this lesson, you will be able to configure WFQ to manage congestion, given a network with suboptimal QoS performance. This ability includes being able to meet these objectives:

- Describe the FIFO queuing mechanism
- Give a detailed explanation of WFQ
- Identify the parameters on which WFQ can classify traffic
- Explain the insertion and drop policy used by WFQ
- Explain how finish time is calculated based on weight and how it is used in the operation of WFQ
- Describe the benefits and drawbacks of using WFQ to implement QoS
- Identify the Cisco IOS commands required to configure WFQ on a Cisco router
- Identify the Cisco IOS commands required to monitor WFQ on a Cisco router

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

FIFO Queuing

This topic describes the FIFO queuing mechanism.

FIFO queuing has no classification because all packets belong to the same class. Packets are dropped when the output queue is full (tail drop). The scheduler services packets in the order they arrive.

The software FIFO queue is basically an extension of the hardware FIFO queue.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

FIFO Queuing (Cont.)

+ **Benefits**
  - **Simple and fast (one single queue with a simple scheduling mechanism)**
  - **Supported on all platforms**
  - **Supported in all switching paths**
  - **Supported in all IOS versions**
– **Drawbacks**
  - **Causes starvation (aggressive flows can monopolize links)**
  - **Causes jitter (bursts or packet trains temporarily fill the queue)**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-3

Although FIFO queuing might be regarded as the fairest queuing mechanism, it has these drawbacks:

- FIFO is extremely unfair when an aggressive flow contests with a fragile flow. Aggressive flows send a large number of packets, many of which are dropped. Fragile flows send a modest number of packets and most of them are dropped because the queue is always full due to the aggressive flow. This type of behavior is called starvation.
- Short or long bursts cause a FIFO queue to fill. Packets entering an almost-full queue have to wait a long time before they can be transmitted. At another time, the queue might be empty, causing packets of the same flow to experience almost no delay. Variation in delay is called jitter.

Despite the drawbacks, FIFO is still the most-used queuing mechanism because of these benefits:

- FIFO is simple and fast. Most high-end routers with fast interfaces are not really challenged by the drawbacks. Furthermore, routers are not capable of complex classification and scheduling when they have to process a large number of packets per second. FIFO is, therefore, the most suitable queuing mechanism on high-end router platforms.
- FIFO is supported on all platforms.
- FIFO queuing is supported in all versions of Cisco IOS.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Weighted Fair Queuing

This topic describes WFQ.

Weighted Fair Queuing

- **A queuing algorithm should share the bandwidth fairly among flows by:**
  - **Reducing response time for interactive flows by scheduling them to the front of the queue**
  - **Preventing high-volume conversations from monopolizing an interface**
- **In the WFQ implementation, messages are sorted into conversations (flows) and transmitted by the order of the last bit crossing its channel.**
- **Unfairness is reinstated by introducing weight to give proportionately more bandwidth to flows with higher IP precedence (lower weight).**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-4

WFQ was introduced as a solution to these problems of the queuing mechanisms:

- FIFO queuing causes starvation, delay, and jitter.
- Priority queuing (PQ) causes starvation of other lower priority classes and suffers from all FIFO problems within each of the four queues that it uses for prioritization.
- Custom queuing (CQ) causes long delays and also suffers from all FIFO problems within each of the 16 queues that it uses for traffic classification.

The idea of WFQ is as follows:

- To have a dedicated queue for each flow (no starvation, delay, or jitter within the queue)
- To allocate bandwidth fairly and accurately among all flows (minimum scheduling delay, guaranteed service)
- To use IP precedence as weight when allocating bandwidth

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ uses automatic classification. Manually defined classes are not supported.

WFQ dropping is not a simple tail drop. WFQ drops packets of the most aggressive flows.

WFQ scheduler is a simulation of a time-division multiplexing (TDM) system. The bandwidth is fairly distributed to all active flows.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ Implementations

- **Implementation parameters**
  - **Queuing platform: central CPU or VIP**
  - **Classification mechanism**
  - **Weighted fairness**
- **Modified tail drop within each queue**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-6

WFQ is supported on most Cisco routers, as well as on Versatile Interface Processors (VIPs). The implementation of WFQ on the VIP is distinct in these ways:

- Classification identifies a flow and assigns a queue to the flow.
- Weight is used for scheduling to give proportionately more bandwidth to flows with a higher IP precedence.
- The tail-dropping scheme is improved to drop packets of the most aggressive flows.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ Classification

This topic describes the parameters on which WFQ can classify traffic.

WFQ classification has to identify individual flows. (The term “conversation” is also used to signify flow) A flow is identified based on the following information taken from the IP header and the TCP or User Datagram Protocol (UDP) headers:

- Source IP address
- Destination IP address
- Protocol number (identifying TCP or UDP)
- Type of service field
- Source TCP or UDP port number
- Destination TCP or UDP port number

These parameters are usually fixed for a single flow, although there are some exceptions:

- A QoS design can mark packets with different IP precedence values even if they belong to the same flow. You should avoid this when using WFQ.
- Some applications change port numbers (for example, TFTP).

If packets of the same flow do not have the same parameters (for example, a different ToS field), the packets can end up in different queues and reordering can occur.

The parameters are used as input for a hash algorithm that produces a fixed-length number that is used as the index of the queue.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ Classification (Cont.)

- A fixed number of per-flow queues is configured.
- A hash function is used to translate flow parameters into a queue number.
- System packets (8 queues) and RSVP flows (if configured) are mapped into separate queues.
- Two or more flows could map into the same queue, resulting in lower per-flow bandwidth.
- Important: the number of queues configured has to be larger than the expected number of flows.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-8

WFQ uses a fixed number of queues. The hash function is used to assign a queue to a flow. There are eight additional queues for system packets and optionally up to 1000 queues for Resource Reservation Protocol (RSVP) flows. The number of dynamic queues WFQ uses by default is based on the interface bandwidth. Using the default interface bandwidth, WFQ uses 256 dynamic queues. The number of queues can be configured in the range between 16 and 4096 (the number must be a power of 2). The default number of dynamic queues for different interface bandwidths is shown in the table.

Bandwidth Range	Number of Dynamic Queues
Less than or equal to 64 kbps	16
More than 64 kbps and less than or equal to 128 kbps	32
More than 128 kbps and less than or equal to 256 kbps	64
More than 256 kbps and less than or equal to 512 kbps	128
More than 512 kbps	256

If there are a large number of concurrent flows, it is likely that two flows could end up in the same queue. You should have several times as many queues as there are flows (on average). This may not be possible in larger environments where the number of concurrent flows is in the thousands.

The probability of two flows ending up in the same flow could be calculated using the following formula:

$$P = 1 - \frac{Queues!}{Queues^{Flows} \cdot (Queues - Flows)!}$$

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The table lists the probability values for three sizes of the WFQ system (64, 128, and 256 queues), with the number of concurrent flows from 5 to 40.

Flows	64 Queues	128 Queues	256 Queues
5	15%	8%	4%
10	52%	30%	16%
15	83%	57%	34%
20	96%	79%	53%
25	100%	92%	70%
30	100%	98%	83%
35	100%	99%	91%
40	100%	100%	96%

The following is the sample calculation of the probability value for 5 flows and 64 queues:

```
Flows: 5
Queues: 64

Probability = 1 - ((64!) / ((64^5) * (59!)))
            = 1 - ((64 * 63 * 62 * 61 * 60) / (64 * 64 * 64 *
64 * 64)
            = 1 - 0.852105618
            = 0.147894382 or 14.7% (15% rounded off)
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ Insertion and Drop Policy

This topic describes the WFQ insertion and drop policy.

WFQ Insertion and Drop Policy

- **WFQ has two modes of dropping:**
  - **Early dropping when the congestion discard threshold is reached**
  - **Aggressive dropping when the hold-queue out limit is reached**
- **WFQ always drops packets of the most aggressive flow**
- **Drop mechanism exceptions**
  - **Packet classified into an empty sub-queue is never dropped**
  - **The packet precedence has no effect on the dropping scheme**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-9

WFQ uses these two parameters that affect the dropping of packets:

- The congestive discard threshold (CDT) is used to start dropping packets of the most aggressive flow, even before the hold-queue limit is reached.
- The hold-queue out limit (HQO) defines the maximum number of packets that can be in the WFQ system at any time.

There are two exceptions to the WFQ insertion and drop policy, as follows:

- If the WFQ system is above the CDT limit, the packet is still enqueued if the per-flow queue is empty.
- The dropping strategy is not directly influenced by IP precedence.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The figure illustrates the dropping scheme of WFQ. The process can be organized into these steps:

Step 1 Drop the new packet if the WFQ system is full (hold-queue limit reached) and the new packet has the worst finish time (the last in the entire system).

Step 2 Drop the packet with the worst finish time in the WFQ system if the system is full. Enqueue the new packet.

Step 3 When the WFQ system is above the CDT limit, a new packet is dropped if that packet is the last in the WFQ system, even though the WFQ system is still within the hold-queue limit.

Step 4 When the WFQ system is above the CDT limit, and if a new packet would not be the last in the WFQ system, the new packet can be enqueued and no other packet is dropped.

When the WFQ system is below both the HQO and the CDT limit, newly arriving packets are enqueued normally.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ Scheduling

This topic describes how finish time is calculated based on weight and how it is used in the operation of WFQ.

The length of queues (for scheduling purposes) is not in packets but in the time it would take to transmit all the packets in the queue. The end result is that WFQ adapts to the number of active flows (queues) and allocates equal amounts of bandwidth to each flow (queue).

The side effect is that flows with small packets (usually interactive flows) get much better service because they do not need a lot of bandwidth. They need low delay, however, which they get because small packets have a low finish time.

The figure illustrates how two queues (queue A and queue B) are contesting for link bandwidth. For this example, assume that the time units are in ms and time T (value 0 is used in the figure) is the starting point.

Queue A receives packets in the following order and times:

- Packet A1 arrives at time T + 0 ms and requires 100 ms to be transmitted.
- Packet A2 arrives at time T + 60 ms (the input interface is obviously faster than the output interface because the arrival time of packet A2 is before the finish time of packet A1) and requires 20 ms to be transmitted.
- Packet A3 arrives at time T + 70 ms (the input interface is obviously much faster than the output interface) and requires 10 ms to be transmitted.

Queue B receives packets in the following order and times:

- Packet B1 arrives at time T + 50 ms and requires 300 ms to be transmitted.
- Packet B2 arrives at time T + 100 ms and also requires 300 ms to be transmitted.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The finish times of packets in Queue A are as follows:

- Packet A1 has a finish time which is the sum of the current time (because the queue was empty at the time of arrival) and the time it takes to transmit this packet (100 ms): FT_{A1} = 0 ms + 100 ms = 100 ms.
- Packet A2 has a finish time which is the sum of the finish time of the last packet in queue A (packet A1) and the time it would take to transmit this packet (20 ms): FT_{A2} = 100 ms + 20 ms = 120 ms.
- Packet A3 has a finish time which is the sum of the finish time of the last packet in queue A (packet A2) and the time it would take to transmit this packet (20 ms): FT_{A3} = 120 ms + 10 ms = 130 ms.

The finish times for the packets in queue B are as follows:

- Packet B1 has a finish time which is the sum of the current time (because the queue was empty at the time of arrival) and the time it takes to transmit this packet (300 ms): FT_{B1} = 50 ms + 300 ms = 350 ms.
- Packet B2 has a finish time which is the sum of the finish time of the last packet in queue B (packet B1) and the time it would take to transmit this packet (300ms): FT_{B2} = 350 ms + 300 ms = 650 ms.

The packets are scheduled into the hardware queue (or the TxQ) in the ascending order of finish times:

1. A1 (100 ms)
2. A2 (120 ms)
3. A3 (130 ms)
4. B1 (350 ms)
5. B2 (650 ms)

Note WFQ prevents reordering of packets within a single flow (conversation). Small packets are automatically preferred over large packets.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The figure introduces the weight into the finish time calculation. The time it takes to transmit the packet is divided by IP Precedence increased by one (to prevent division by zero).

The WFQ implementation in Cisco routers is optimized in these ways:

- The real time it takes to transmit the packet is not relevant. The packet size can be used instead because it is proportional to the transmit time.
- The packet size is not divided by IP precedence (division is a CPU-intensive operation). Instead, the size is multiplied by a fixed value (one multiplication value for each IP Precedence value).

Packets with IP Precedence one appear half the size they really are. The result is that these packets receive twice as much bandwidth as packets with IP Precedence zero.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Finish Time Calculation with Weights

- **Finish time is adjusted based on IP precedence of the packet.**

If Flow F Active,
Then $FT(P_{k+1}) = FT(P_k) + Size(P_{k+1})/(IPPrec+1)$
Otherwise $FT(P_0) = Now + Size(P_0)/(IPPrec+1)$

- **IOS implementation scales the finish time to allow integer arithmetic.**

If Flow F Active,
Then $FT(P_{k+1}) = FT(P_k) + Size(P_{k+1})*32384/(IPPrec+1)$
Otherwise $FT(P_0) = Now + Size(P_0)*32384 /(IPPrec+1)$

- **RSVP packets and high-priority internal packets have special weights (4 and 128).**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-13

The first formula in the figure is the first optimization in which the finish time is really the sum of packet sizes divided by an increased IP precedence value.

The second formula shows further optimization in which, instead of dividing, the packet size is multiplied by 32,384 / (IP Precedence + 1). A number for each different IP precedence value is stored in a table and therefore does not have to be calculated for each packet.

Packets belonging to RSVP flows and system packets have special low weights that guarantee them more bandwidth.

Note Cisco IOS versions before Release 12.0(5)T use a new formula in which the weight is calculated on this formula: Weight = 4096 / (IP Precedence +1).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

IP Precedence to Weight Mapping

IP Precedence	Weight
0	32384
1	16192
2	10794
3	8096
4	6476
5	5397
6	4626
7	4048
32 (virtual IP precedence)	128
1024 (virtual IP precedence)	4 (RSVP)

01FQ_124

- **RSVP packets and high-priority internal packets have special weights (4 and 128).**
- **Lower weight makes packets appear smaller (preferred).**
- **These numbers are subject to change.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-14

The figure shows the mapping between IP precedence values and WFQ weights.

Note These figures are subject to change. Refer to the Cisco IOS documentation for the latest information on WFQ weights.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

This case study describes how packets are dropped in different situations.

WFQ Case Study

- **WFQ system can hold a maximum of ten packets (hold-queue limit).**
- **Early dropping (of aggressive flows) should start when there are eight packets (congestive discard threshold) in the WFQ system.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-15

The WFQ system was reduced to a modest hold-queue limit of ten and a congestive discard threshold of eight.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ Case Study
Interface Congestion

- **HQO (hold-queue out limit) is the maximum number of packets that the WFQ system can hold and HQO = 10.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-16

There are already ten packets in the WFQ system. The new packet would be the 11th and also the last in the entire WFQ system.

WFQ Case Study
Interface Congestion (Cont.)

- **HQO is the maximum number of packets that the WFQ system can hold and HQO = 10.**
- **Absolute maximum (HQO=10) exceeded, new packet is the last in the TDM system and is dropped.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-17

The new packet is dropped.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ Case Study
Flow Congestion

- **Early dropping (of aggressive flows) should start when there are eight packets (congestive discard threshold) in the WFQ system.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-18

The example illustrates how WFQ can drop packets even if the WFQ system is still within the hold-queue limit. The system, however, is above the CDT limit. In this case, a packet can be dropped if it is the last in the system.

WFQ Case Study
Flow Congestion (Cont.)

- **Early dropping (of aggressive flows) should start when there are eight packets (congestive discard threshold) in the WFQ system.**
- **CDT exceeded (CDT=8), new packet would be the last in the TDM system and is dropped.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-19

The new packet would be the last in the TDM system and is dropped.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Benefits and Drawbacks of WFQ

This topic describes the benefits and drawbacks of using WFQ to implement QoS.

Benefits and Drawbacks of WFQ

+ **Benefits**
  - **Simple configuration (classification does not have to be configured)**
  - **Guarantees throughput to all flows**
  - **Drops packets of most aggressive flows**
  - **Supported on most platforms**
  - **Supported in all IOS versions**

– **Drawbacks**
  - **Multiple flows can end up in one queue**
  - **Does not support the configuration of classification**
  - **Cannot provide fixed bandwidth guarantees**
  - **Complex classification and scheduling mechanisms**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-20

These are the main benefits of WFQ:

- Simple configuration (no manual classification is necessary)
- Guarantees throughput to all flows
- Drops packets of the most aggressive flows
- Supported on most platforms
- Supported in all Cisco IOS versions

These are the main drawbacks of WFQ:

- Not always possible to have one flow per queue
- Does not allow manual classification
- Cannot provide fixed guarantees
- Complex classification and scheduling mechanisms

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring WFQ

This topic describes the Cisco IOS commands required to configure WFQ on a Cisco router.

Configuring WFQ

```
router(config-intf)#
fair-queue [cdt [dynamic-queues [reservable-queues]]]
```

- **CDT**
  - **Number of messages allowed in the WFQ system before the router starts dropping new packets for the longest queue.**
  - **The value can be in the range from 1 to 4096 (default is 64)**
- **dynamic-queues**
  - **Number of dynamic queues used for best-effort conversations (values are: 16, 32, 64, 128, 256, 512, 1024, 2048, and 4096)**
- **reservable-queues**
  - **Number of reservable queues used for reserved conversations in the range 0 to 1000 (used for interfaces configured for features such as RSVP - the default is 0)**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-21

WFQ is automatically enabled on all interfaces that have a default bandwidth of less than 2 Mbps. The **fair-queue** command is used to enable WFQ on interfaces where it is not enabled by default or was previously disabled.

fair-queue [*congestive-discard-threshold* [*dynamic-queues* [*reservable-queues*]]]

Syntax Description

Parameter	Description
congestive-discard-threshold	(Optional) Number of messages allowed in the WFQ system. The default is 64 messages, and a new threshold must be a power of 2 in the range from 16 to 4096. When a conversation reaches this threshold, new message packets are discarded.
dynamic-queues	(Optional) Number of dynamic queues used for best-effort conversations. Values are 16, 32, 64, 128, 256, 512, 1024, 2048, and 4096.
reservable-queues	(Optional) Number of reservable queues used for reserved conversations in the range 0 to 1000. The default is 0. Reservable queues are used for interfaces configured for features such as RSVP.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Additional WFQ Configuration Parameters

```
router(config-if)#
hold-queue max-limit out
```

- **Specifies the maximum number of packets that can be in all output queues on the interface at any time.**
- **The default value for WFQ is 1000.**
- **Under special circumstances, WFQ can consume a lot of buffers, which may require lowering this limit.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-22

The same **hold-queue** command that can be used with FIFO queuing can also be used with WFQ. The default hold-queue limit with WFQ is 1000 packets.

The WFQ system will generally never reach the hold-queue limit because the CDT limit starts dropping packets of aggressive flows. Under special circumstances, it would be possible to fill the WFQ system. For example, a denial-of-service attack that floods the interface with a large number of packets (each different) could fill all queues at the same rate.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WFQ Configuration Defaults

- **Fair queuing is enabled by default:**
  - **On physical interfaces whose bandwidth is less than or equal to 2.048 Mbps**
  - **On interfaces configured for Multilink PPP**
- **Fair queuing is disabled:**
  - **If you enable the autonomous or silicon switching engine mechanisms**
  - **For any sequenced encapsulation: X.25, SDLC, LAPB, reliable PPP**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-23

The figure explains the default behavior of WFQ. WFQ is automatically enabled on all interfaces slower than 2 Mbps. WFQ is also required on interfaces using Multilink PPP (MLP).

WFQ cannot be used if the switching path does not support WFQ, or if reordering of frames is not allowed because of sequence numbering of Layer 2 frames. WFQ is disabled for any sequenced encapsulation, such as X.25, Synchronous Data Link Control (SDLC), Link Access Procedure, Balanced (LAPB), and reliable PPP.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring WFQ

This topic describes the Cisco IOS commands required to monitor WFQ on a Cisco router.

Monitoring WFQ

```
router>
show interface interface
```

- **Displays interface delays including the activated queuing mechanism with the summary information**

```
Router>show interface serial 1/0
  Hardware is M4T
  Internet address is 20.0.0.1/8
  MTU 1500 bytes, BW 19 Kbit, DLY 20000 usec, rely 255/255, load
147/255
  Encapsulation HDLC, crc 16, loopback not set
  Keepalive set (10 sec)
  Last input 00:00:00, output 00:00:00, output hang never
  Last clearing of "show interface" counters never
  Input queue: 0/75/0 (size/max/drops); Total output drops: 0
  Queueing strategy: weighted fair
  Output queue: 0/1000/64/0 (size/max total/threshold/drops)
     Conversations  0/4/256 (active/max active/max total)
     Reserved Conversations 0/0 (allocated/max allocated)
  5 minute input rate 18000 bits/sec, 8 packets/sec
  5 minute output rate 11000 bits/sec, 9 packets/sec
… rest deleted ...
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-24

Use the same **show** commands as with other queuing mechanisms:

- **show interface**
- **show queue**
- **show queueing**

Use the **show interface** command to determine the queuing strategy. The summary statistics are also displayed.

The sample output in the figure shows that there are currently no packets in the WFQ system. The system allows up to 1000 packets (hold-queue limit) with CDT 64. WFQ is using 256 queues. The maximum number of concurrent conversations (active queues) is four.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring WFQ (Cont.)

```
router>
show queue interface-name interface-number
```

- **Displays detailed information about the WFQ system of the selected interface**

```
Router>show queue serial 1/0
  Input queue: 0/75/0 (size/max/drops); Total output drops: 0
  Queueing strategy: weighted fair
  Output queue: 2/1000/64/0 (size/max total/threshold/drops)
     Conversations  2/4/256 (active/max active/max total)
     Reserved Conversations 0/0 (allocated/max allocated)

  (depth/weight/discards/tail drops/interleaves) 1/4096/0/0/0
  Conversation 124, linktype: ip, length: 580
  source: 193.77.3.244, destination: 20.0.0.2, id: 0x0166, ttl: 254,
  TOS: 0 prot: 6, source port 23, destination port 11033

  (depth/weight/discards/tail drops/interleaves) 1/4096/0/0/0
  Conversation 127, linktype: ip, length: 585
  source: 193.77.4.111 destination: 40.0.0.2, id: 0x020D, ttl: 252,
  TOS: 0 prot: 6, source port 23, destination port 11013
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-25

The **show queue** command is used to display the contents of packets inside a queue for a particular interface, including flow (conversation) statistics, as follows:

- Queue depth is the number of packets in the queue.
- Weight is 4096 / (IP Precedence + 1), or 32,384 / (IP Precedence + 1), depending on the Cisco IOS version.
- In the command output, discards are used to represent the number of drops due to the CDT limit.
- In the command output, tail drops are used to represent the number of drops due to the hold-queue limit.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **The software FIFO queue is basically an extension of the hardware FIFO queue.**
- **WFQ was developed to overcome the limitations of the more basic queuing methods. Traffic is sorted into flows and transmitted by the order of the last bit crossing its channel.**
- **WFQ classification uses as parameters: source and destination IP addresses, source and destination TCP or UDP ports, transport protocol, and ToS field.**
- **With WFQ, the CDT is used to start dropping packets of the most aggressive flow, even before the hold-queue limit is reached, and the hold-queue out limit defines the total maximum number of packets that can be in the WFQ system at any time.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-26

Summary (Cont.)

- **Finish time, the time it takes to transmit the packet, is divided by IP precedence increased by one (to prevent division by zero).**
- **WFQ benefits: Simple configuration, drops packets of the most aggressive flows. WFQ drawbacks: Not always possible to have one flow per queue, does not allow manual classification, and cannot provide fixed guarantees.**
- **WFQ is automatically enabled on all interfaces that have a default bandwidth of less than 2 Mbps. The** fair-queue **command is used to enable WFQ on interfaces where it is not enabled by default or was previously disabled.**
- **The same** show **commands can be used as with other queuing mechanisms:** show interface**,** show queue, **and** show queuing**.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-27

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 4

Configuring CBWFQ and LLQ

Overview

Class-based weighted fair queuing (CBWFQ) extends the standard weighted fair queuing (WFQ) functionality to provide support for user-defined traffic classes. With CBWFQ, you define traffic classes based on match criteria including Network-Based Application Recognition (NBAR), access control lists (ACLs), and input interfaces. Packets satisfying the match criteria for a class constitute the traffic for that class. A queue is reserved for each class, and traffic belonging to a class is directed to the queue for that class. low-latency queuing (LLQ) brings strict priority queuing to CBWFQ. Strict priority queuing allows delay-sensitive data such as voice to be dequeued and sent first (before packets in other queues are dequeued), giving delay-sensitive data preferential treatment over other traffic. This lesson describes how to configure CBWFQ and LLQ.

Objectives

Upon completing this lesson, you will be able to configure CBWFQ and LLQ to manage congestion. This ability includes being able to meet these objectives:

- Explain how basic queuing mechanisms can be used to build advanced queuing mechanisms
- Explain the purpose and features of CBWFQ
- Describe CBWFQ features and explain how CBWFQ works using a block diagram
- Describe the benefits of CBWFQ
- Identify the Cisco IOS commands required to configure and monitor CBWFQ on a Cisco router
- Explain the purpose and features of LLQ
- Explain how LLQ works using a block diagram and identify situations in which LLQ is most appropriate for providing QoS
- Describe the benefits of LLQ
- Identify the Cisco IOS commands required to configure and monitor LLQ on a Cisco router

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CBWFQ and LLQ

This topic describes how you can use basic queuing mechanisms to build more advanced queuing mechanisms.

Neither the basic queuing methods nor the more advanced WFQ completely solve quality of service (QoS) problems resulting from converged network traffic. These problems remain:

- If only priority queuing (PQ) is used for a voice-enabled network, voice will get the priority needed, but data traffic will likely starve out at some point.
- If only custom queuing (CQ) is used for a voice-enabled network, data traffic will be assured of some bandwidth, but voice traffic will suffer excessive delays.
- If WFQ is used, voice still experiences delay, even when treated “fairly” by WFQ and no bandwidth guarantees can be made for data applications.
- All of the classification, marking, and queuing mechanisms are complicated to use and time-consuming when applied on an interface-by-interface basis.

Newer queuing mechanisms take the best aspects of existing queuing methods and apply them to give voice the priority it requires while still ensuring that data is serviced efficiently on a class basis.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based Weighted Fair Queuing

This topic describes the purpose and features of CBWFQ.

Class-Based Weighted Fair Queuing

- **CBWFQ is a mechanism that is used to guarantee bandwidth to classes.**
- **CBWFQ extends the standard WFQ functionality to provide support for user-defined traffic classes.**
  - **Classes are based on user-defined match criteria.**
  - **Packets satisfying the match criteria for a class constitute the traffic for that class.**
- **A queue is reserved for each class, and traffic belonging to a class is directed to that class queue.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-3

CBWFQ extends the standard WFQ functionality to provide support for user-defined traffic classes. With CBWFQ, you define the traffic classes based on match criteria, including protocols, ACLs, and input interfaces. Packets satisfying the match criteria for a class constitute the traffic for that class. A queue is reserved for each class, and traffic belonging to a class is directed to that class queue.

After a class has been defined according to its match criteria, you can assign characteristics to it. To characterize a class, you assign it bandwidth, weight, and maximum packet limit. The bandwidth assigned to a class is the minimum bandwidth delivered to the class during congestion.

To characterize a class, you also specify the queue limit for that class, which is the maximum number of packets allowed to accumulate in the class queue. Packets belonging to a class are subject to the bandwidth and queue limits that characterize the class. After a queue has reached its configured queue limit, enqueuing of additional packets to the class causes tail drop or random packet drop to take effect, depending on how the class policy is configured.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CBWFQ Architecture

This topic describes the features of CBWFQ and how CBWFQ works.

CBWFQ supports multiple class maps (the number depends upon the platform) to classify traffic into its corresponding FIFO queues. Tail drop is the default dropping scheme of CBWFQ. You can use weighted random early detection (WRED) in combination with CBWFQ to prevent congestion of a class.

The CBWFQ scheduler is used to guarantee bandwidth that is based on the configured weights.

Note Currently, except for the Cisco 7500 series router platform, all traffic classes (default traffic class excluded) only support FIFO queuing within the class. On all platforms, the default traffic class can support either FIFO or WFQ within the class. Check Cisco.com for the latest information on WFQ support within each traffic class.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CBWFQ Architecture: Classification

- **Classification uses class maps.**
- **Availability of certain classification options depends on the Cisco IOS version.**
- **Some classification options depend on type of interface and encapsulation where service policy is used.**
- **For example:**
  - **Matching on Frame Relay discard eligible bits can only be used on interfaces with Frame Relay encapsulation.**
  - **Matching on MPLS experimental bits has no effect if MPLS is not enabled.**
  - **Matching on ISL priority bits has no effect if ISL is not used.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-9

You can use any classification option depending on its availability in the Cisco IOS version and the support on the selected interface and encapsulation. These examples illustrate some of the limitations regarding classification options:

- Matching on Frame Relay discard eligible bits can only be used on interfaces with Frame Relay encapsulation.
- Matching on Multiprotocol Label Switching (MPLS) Experimental bits has no effect if MPLS is not enabled.
- Matching on Inter-Switch Link (ISL) priority bits has no effect if ISL is not used.

It is important to note that CBWFQ is configured using Modular QoS command-line interface (CLI), or MQC.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CBWFQ Architecture: Insertion Policy

- **Each queue has a maximum number of packets that it can hold (queue size).**
- **The maximum queue size is platform-dependent.**
- **After a packet is classified to one of the queues, the router will enqueue the packet if the queue limit has not been reached (tail drop within each class).**
- **WRED can be used in combination with CBWFQ to prevent congestion of the class.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-10

CBWFQ reserves multiple FIFO queues in the WFQ system. The default queue limit is 64 (tail drop) and can be configured with WRED (weighted random drop).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CBWFQ Architecture: Scheduling

- **CBWFQ guarantees bandwidth according to weights assigned to traffic classes.**
- **Weights can be defined by specifying:**
  - **Bandwidth (in kbps)**
  - **Percentage of bandwidth (percentage of available interface bandwidth)**
  - **Percentage of remaining available bandwidth**
- **One service policy can not have mixed types of weights.**
- **The** show interface **command can be used to display the available bandwidth.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-11

You can configure bandwidth guarantees by using one of these commands:

- The **bandwidth** command allocates a fixed amount of bandwidth by specifying the amount in kbps. The reserved bandwidth is subtracted from the available bandwidth of the interface where the service policy is used. The allocated bandwidth must also be within the configured reservable limit (75 percent by default).
- You can use the **bandwidth percent** command to allocate a percentage of the default or available bandwidth of an interface. The default bandwidth usually equals the maximum speed of an interface. Sometimes it actually reflects the real speed of an interface (for example, Ethernet or FastEthernet). The default value can be replaced by using the **bandwidth** interface command. It is recommended that the bandwidth reflect the real speed of the link. The allocated bandwidth is subtracted from the available bandwidth of the interface in which the service policy is used.
- You can use the **bandwidth remaining percent** command to allocate a portion of the unallocated bandwidth. The bandwidth is *not* subtracted from the available bandwidth of the interface where the service policy is used.

A single service policy cannot mix the fixed bandwidth (in kbps) and bandwidth percent commands (except with strict priority queues).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CBWFQ Architecture: Available Bandwidth

- **Available bandwidth is calculated according to the following formula:**

Configured using the interface bandwidth **command**

Configured using the interface max-reserved-bandwidth **command. 75% is the percent of bandwidth which can be used by CBWFQ and LLQ.** max-reserved-bandwidth **can be used to override the 75% limit, but this is NOT recommended.**

017Q_160

$$Bandwidth_{avail} = Bandwidth * MaxReservable - SUM(all\ fixed\ guarantees)$$

Sum of all fixed guarantees using CBWFQ, CB-LLQ, IP RTP prioritization

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-12

The available bandwidth displayed by the **show interface** command is calculated by subtracting all fixed bandwidth reservations from 75 percent of the configured bandwidth of an interface.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CBWFQ Architecture: 75 Percent Rule

- **Add**
  - **Class bandwidths**
  - **RSVP maximum reserved bandwidth**
- **Result must be less than or equal to 75% of interface bandwidth (or Frame Relay, DLCI, CIR)**
  - **Leaves headroom for overhead traffic such as Layer 2 keepalives and bandwidth for the class default traffic**
- **The 75% rule is a conservative rule**
- max-reserved-bandwidth **command overrides 75% limit, but seldom recommended**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-13

Properly provisioning the network bandwidth is a major component of successful network design. You can calculate the required bandwidth by adding the bandwidth requirements for each major application (for example, voice, video, and data). The resulting sum represents the minimum bandwidth requirement for any given link, and it should not exceed 75 percent of the total available bandwidth for the link (or Frame Relay or ATM PVC committed information Rate [CIR]). This 75 percent rule assumes that some bandwidth is required for overhead traffic, such as layer 2 keepalive messages, as well as for the class default traffic.

The total amount of bandwidth allocated for all classes included in a policy map should not exceed 75 percent of the available bandwidth on the interface. The **max-reserved bandwidth** command overrides the 75 percent limitation, but overriding is recommended only for the most knowledgeable network administrators who have access to precise figures for available, used, and required bandwidth, including control plane traffic bandwidth requirements. If all of the bandwidth is not allocated, the remaining bandwidth is proportionally allocated among the classes based on the configured bandwidth of the classes.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CBWFQ Benefits

This topic describes the benefits of CBWFQ.

CBWFQ Benefits

+ Benefits
  - Minimum bandwidth allocation
  - Finer granularity and scalability
  - MQC interface easy to use
  - Maximizes transport of priority traffic
  - Weights guarantee minimum bandwidth
  - Unused capacity shared among the other classes
  - Queues separately configured for QoS
– Drawbacks
  - Voice traffic can still suffer unacceptable delay

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-14

CBWFQ allows you to define traffic classes based on custom-defined match criteria such as ACLs, input interfaces, protocol, and QoS label. For example, a class might consist of a team working on a certain project, or a class can be created for the important mission-critical applications such as enterprise resource planning (ERP) applications. When the traffic classes have been defined, they can be assigned a bandwidth, queue limit, or drop policy such as WRED. These are additional benefits of CBWFQ:

- **Bandwidth allocation:** CBWFQ allows you to specify the exact amount of bandwidth to be allocated for a specific class of traffic. Accounting for available bandwidth on the interface, you can configure multiple classes (the number depends upon the platform).
- **Finer granularity and scalability:** CBWFQ allows you total flexibility to define a class, based on ACLs and protocols or input interfaces, thereby providing finer granularity.
- **Supported by MQC:** CBWFQ is supported by the easy-to-use MQC.

The CBWFQ feature is supported on all platforms that WFQ is supported on; in other words, the Cisco 7200, 3800, 3800, 1800, 800, and others.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring and Monitoring CBWFQ

This topic describes the Cisco IOS commands that are used to configure and monitor CBWFQ on a Cisco router.

Configuring CBWFQ

```
router(config-pmap-c)#
bandwidth bandwidth
```

- **Allocates a fixed amount of bandwidth to a class**
- **Sets the value in kbps**

```
router(config-pmap-c)#
bandwidth percent percent
```

- **Allocates a percentage of bandwidth to a class.**
- **The configured (or default) interface bandwidth is used to calculate the guaranteed bandwidth.**

```
router(config-pmap-c)#
bandwidth remaining percent percent
```

- **Allocates a percentage of available bandwidth to a class**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-15

The **bandwidth** policy-map class configuration command is used to specify or modify the bandwidth allocated for a class belonging to a policy map.

All classes belonging to one policy map should use the same type of fixed bandwidth guarantee, in kbps, percentage of interface bandwidth, or percentage of available bandwidth.

Configuring bandwidth in percentages is most useful when the underlying link bandwidth is unknown or the relative class bandwidth distributions are known.

bandwidth {*bandwidth-kbps* | **remaining percent** *percentage* | **percent** *percentage*}

Syntax Description

Parameter	Description
bandwidth-kbps	Amount of bandwidth, in kbps, to be assigned to the class.
remaining percent *percentage*	Amount of guaranteed bandwidth, based on a relative percent of available bandwidth. The percentage can be a number from 1 to 100.
percent *percentage*	Amount of guaranteed bandwidth, based on an absolute percent of available bandwidth. The percentage can be a number from 1 to 100. (By default, only 75 percent can be reserved.)

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

These restrictions apply to the **bandwidth** command:

- If the **percent** keyword is used, the sum of the class bandwidth percentages cannot exceed 100 percent.
- The amount of bandwidth configured should be large enough to also accommodate Layer 2 overhead.
- A policy map can have all the class bandwidths specified in kbps or in percentages, but not a mix of both. However, the unit for the **priority** command in the priority class can be different from the bandwidth unit of the low-priority class.

CBWFQ is not supported on sub-interfaces. If a service policy is applied to a sub-interface in which the service policy directly references a policy map that has a bandwidth statement in it, this error message will be displayed:

```
router(config-subif)#service-policy output cbwfq
CBWFQ : NOT supported on subinterfaces
```

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring CBWFQ (Cont.)

```
router(config-pmap-c)#
queue-limit queue-limit
```

- **Sets the maximum number of packets this queue can hold**
- **The default maximum is 64**

```
router(config-pmap-c)#
fair-queue [congestive-discard-threshold [dynamic-queues]]
```

- **The "class-default" class can be configured to use WFQ.**
- **Number of messages allowed in each queue. The default is 64 messages, and a new threshold must be a power of 2 in the range from 16 to 4096.**
- **The number of dynamic queues is a power of 2 number in the range from 16 to 4096, specifying the number of dynamic queues**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-16

You can change the default queue limit of 64 packets by using the **queue-limit** command. It is recommended that you not change the default value. The default class can be selected by specifying the **class-default** name of the class. The default class supports two types of queuing: a FIFO queue (default) or a flow-based WFQ system. Both types can be combined with WRED. A FIFO queue can also get a bandwidth guarantee.

Example: Configuration of FIFO Queuing

This example shows the configuration of FIFO queuing within the default class. The default class is guaranteed 1 Mbps of bandwidth and the maximum queue size is limited to 40 packets.

```
policy-map A
 class class-default
  bandwidth 1000
  queue-limit 40
```

Example: Configuration of WFQ Queuing

This example shows the configuration of WFQ queuing within the default class. The number of dynamic queues is set to 1024 and the discard threshold is set to 32.

```
policy-map A
 class class-default
  fair-queue 32 1024
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring CBWFQ (Cont.)

```
Router(config)# access-list 101 permit udp host 10.10.10.10 host
10.10.10.20 range 16384 20000
Router(config-if)# access-list 101 permit udp host  10.10.10.10
host 10.10.10.20 range 53000 56000
Router(config)# class-map class1
Router(config-cmap)# match access-group 101
Router(config-cmap)# exit
Router(config-cmap)# class-map class2
Router(config-cmap)# match access-group 102
Router(config-cmap)# exit

Router(config)# policy-map policy1
Router(config-pmap)# class class1
Router(config-pmap-c)# bandwidth 3000
Router(config-pmap-c)# queue-limit 30
Router(config-pmap-c)# exit
Router(config-pmap)# class class2
Router(config-pmap-c)# bandwidth 2000
Router(config-pmap-c)# exit
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-17

The sample configuration shows how CBWFQ is used to guarantee bandwidth to each of the two classes.

Monitoring CBWFQ

```
router>
show policy-map interface [interface]
```

- **Displays parameters and statistics of CBWFQ**

```
router>show policy-map interface
 FastEthernet0/0

  Service-policy output: Policy1

    Class-map: Class1 (match-any)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: any
      Weighted Fair Queueing
        Output Queue: Conversation 265
        Bandwidth remaining 20 (%) Max Threshold 64 (packets)
        (pkts matched/bytes matched) 0/0
        (depth/total drops/no-buffer drops) 0/0/0
    Class-map: class-default (match-any)
      42 packets, 4439 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: any
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-18

The **show policy-map interface** command displays all service policies applied to the interface. The output displays the policing parameters and statistics.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Low-Latency Queuing

This topic describes the purpose and features of LLQ.

Low-Latency Queuing

- **Priority queue added to CBWFQ for real-time traffic**
- **High-priority classes are guaranteed:**
  - **Low-latency propagation of packets**
  - **Bandwidth**
- **High-priority classes are also policed when congestion occurs – they can then not exceed their guaranteed bandwidth**
- **Lower priority classes use CBWFQ**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-19

While WFQ provides a fair share of bandwidth to every flow, and provides fair scheduling of its queues, it cannot provide guaranteed bandwidth and low delay to select applications. For example, voice traffic may still compete with other aggressive flows in the WFQ queuing system because the WFQ system lacks priority scheduling for time-critical traffic classes.

The LLQ feature brings strict priority queuing to CBWFQ. Strict priority queuing allows delay-sensitive data such as voice to be dequeued and sent first (before packets in other queues are dequeued), giving delay-sensitive data preferential treatment over other traffic.

Without LLQ, CBWFQ provides weighted fair queuing based on defined classes with no strict priority queue available for real-time traffic. CBWFQ allows you to define traffic classes and then assign characteristics to that class. For example, you can designate the minimum bandwidth delivered to the class during congestion.

For CBWFQ, the weight for a packet belonging to a specific class is derived from the bandwidth that you assigned to the class when you configured it. Therefore, the bandwidth assigned to the packets of a class determines the order in which packets are sent. All packets are serviced fairly based on weight; no class of packets may be granted strict priority. This scheme poses problems for voice traffic, which is largely intolerant of delay, especially variation in delay. For voice traffic, variations in delay introduce irregularities of transmission that are heard as jitter in the conversation.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The LLQ feature provides strict priority queuing for CBWFQ, reducing jitter in voice conversations. Configured by the **priority** command, LLQ enables use of a single, strict priority queue within CBWFQ at the class level, allowing you to direct traffic belonging to a class to the CBWFQ strict priority queue. To enqueue class traffic to the strict priority queue, you configure the **priority** command for the class after you specify the named class within a policy map. Classes to which the priority command is applied are considered priority classes. Within a policy map, you can give one or more classes priority status. When multiple classes within a single policy map are configured as priority classes, all traffic from these classes is enqueued to the same, single, strict priority queue.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

LLQ Architecture

This topic describes how LLQ works and identifies situations in which LLQ is most appropriate for providing QoS.

When CBWFQ is configured as the queuing system, it creates a number of queues, into which it classifies traffic classes. These queues are then scheduled with a WFQ-like scheduler, which can guarantee bandwidth to each class.

If LLQ is used within the CBWFQ system, it creates an additional priority queue in the WFQ system, which is serviced by a strict priority scheduler. Any class of traffic can therefore be attached to a service policy, which uses priority scheduling, and hence can be prioritized over other classes.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

LLQ Benefits

This topic describes the benefits of LLQ.

LLQ Benefits

Benefits

- **High-priority classes are guaranteed:**
  - **Low-latency propagation of packets**
  - **Bandwidth**
- **Consistent configuration and operation across all media types**
- **Entrance criteria to a class can be defined by an ACL**
  - **Not limited to UDP ports as with IP RTP priority**
  - **Defines trust boundary to ensure simple classification and entry to a queue**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-25

The LLQ priority scheduler guarantees both low-latency propagation of packets and bandwidth to high-priority classes. Low latency is achieved by expediting traffic using a priority scheduler. Bandwidth is also guaranteed by the nature of priority scheduling, but is policed to a user-configurable value. The strict PQ scheme allows delay-sensitive data such as voice to be dequeued and sent first—that is, before packets in other queues are dequeued. Delay-sensitive data is given preferential treatment over other traffic.

The LLQ feature provides strict PQ on ATM virtual circuits (VCs); the IP Real-Time Transport Protocol (RTP) priority feature only allows PQ on interfaces. Because you can configure the priority status for a class within CBWFQ, you are not limited to User Datagram Protocol (UDP) port numbers to stipulate priority flows (which were necessary with IP RTP). Instead, all of the valid match criteria used to specify traffic for a class now applies to priority traffic.

Policing of priority queues also prevents the priority scheduler from monopolizing the CBWFQ scheduler and starving nonpriority classes, as legacy PQ does. By configuring the maximum amount of bandwidth allocated for packets belonging to a class, you can avoid starving nonpriority traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring and Monitoring LLQ

This topic describes the Cisco IOS commands that are used to configure and monitor LLQ on a Cisco router.

Configuring LLQ

```
router(config-pmap-c)#
priority bandwidth [burst]
```

- **Allocates a fixed amount of bandwidth (in kbps) to a class and ensures expedited forwarding.**
- **Traffic exceeding the specified bandwidth is dropped if congestion exists; otherwise, policing is not used.**

```
router(config-pmap-c)#
priority percent percentage [burst]
```

- **Allocates a percentage of configured or default interface bandwidth to a class and ensures expedited forwarding.**
- **Traffic exceeding the specified bandwidth is dropped if congestion exists; otherwise, policy is not used.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-26

When you specify the **priority** command for a class, it takes a bandwidth argument that gives maximum bandwidth in kbps. You use this parameter to specify the maximum amount of bandwidth allocated for packets belonging to the class configured with the **priority** command. The bandwidth parameter both guarantees bandwidth to the priority class and restrains the flow of packets from the priority class.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

priority {*bandwidth-kbps* | **percent** *percentage*} [**burst**]

Parameter	Description
bandwidth-kbps	Guaranteed allowed bandwidth, in kbps, for the priority traffic. The amount of guaranteed bandwidth varies according to the interface and platform in use. Beyond the guaranteed bandwidth, the priority traffic will be dropped in the event of congestion to ensure that the nonpriority traffic is not starved.
percent	Specifies that the amount of guaranteed bandwidth will be specified by the percent of available bandwidth.
percentage	Used in conjunction with the **percent** keyword; specifies the percentage of the total available bandwidth to be set aside for the priority class. The percentage can be a number from 1 to 100. (By default, only 75 percent can be reserved.)
burst	(Optional) Specifies the burst size, in bytes. The range of the burst is 32 to 2,000,000 bytes. The burst size allows temporary bursting of traffic above the maximum limit after a period of inactivity. The default burst value is computed as 200 ms of traffic at the configured bandwidth rate.

In the event of congestion, if the priority class traffic exceeds the bandwidth guarantee, a congestion-aware policer is used to drop the exceeding packets. Voice traffic enqueued to the priority queue is UDP-based and therefore not adaptive to the early packet-drop characteristic of WRED. Because WRED is ineffective, you cannot use the WRED **random-detect** command with the **priority** command. In addition, because policing is used to drop packets and a queue limit is not imposed, the **queue-limit** command cannot be used with the **priority** command.

When congestion occurs, traffic destined for the priority queue is metered to ensure that the bandwidth allocation configured for the class to which the traffic belongs is not exceeded.

Priority traffic metering has these qualities:

- It is much like committed access rate (CAR), except that priority traffic metering is only performed under conditions of congestion. When the device is not congested, the priority-class traffic is allowed to exceed its allocated bandwidth. When the device is congested, the priority-class traffic above the allocated bandwidth is discarded.
- Metering performed on a per-packet basis, and tokens, are replenished as packets are sent. If not enough tokens are available to send the packet, the packet is dropped.
- It restrains priority traffic to its allocated bandwidth to ensure that nonpriority traffic, such as routing packets and other data, is not starved.

With metering, the classes are policed and rate-limited individually. That is, although a single policy map might contain four priority classes, all of which are enqueued in a single priority queue, they are each treated as separate flows with separate bandwidth allocations and constraints.

Note It is important to note that because bandwidth for the priority class is specified as a parameter to the **priority** command, you cannot also configure the **bandwidth** command for a priority class. To do so is a configuration violation that would only introduce confusion in relation to the amount of bandwidth to allocate.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Keep these guidelines in mind when using the **priority** command:

- You account for Layer 2 encapsulations in the amount of bandwidth that you specify with the **priority** command. However, ensure that you configure a bandwidth with room for the cell-tax overhead.
- Use the **priority** command for Voice over IP (VoIP) on serial links and ATM permanent virtual circuits (PVCs).

Example: Calculating LLQ Bandwidth Required for VoIP

The bandwidth consumed by VoIP streams is calculated by adding the packet payload and all the headers, then multiplying that total number by the per-second packet rate.

The example shows how to calculate the VoIP bearer bandwidth requirement for a single VoIP call using a G.711 codec:

```
G.711 = 160 bytes payload size


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


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


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


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


Bandwidth (bits/sec) with Layer 2 overhead


   = 80000 bits/second + L2 overhead bytes/second
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring LLQ (Cont.)

```
class-map voip
 match ip precedence 5
!
class-map mission-critical
 match ip precedence 3 4
!
class-map transactional
 match ip precedence 1 2
!
policy-map Policy1
  class voip
   priority percent 10
  class mission-critical
   bandwidth percent 30
   random-detect
  class transactional
   bandwidth percent 20
   random-detect
  class class-default
   fair-queue
   random-detect
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-27

The figure shows a configuration example in which the VoIP traffic class, classified by the IP Precedence of 5, is queued in a LLQ within the CBWFQ system. The VoIP class receives guaranteed priority scheduling over the other classes but is limited to 10 percent of the bandwidth.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring LLQ

```
router>
show policy-map interface interface
```

- **Displays the packet statistics of all classes that are configured for all service policies either on the specified interface or subinterface**

```
router>show policy-map interface fastethernet 0/0
 FastEthernet0/0

  Service-policy output: LLQ

    Class-map: LLQ (match-any)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: any
      Weighted Fair Queueing
        Strict Priority
        Output Queue: Conversation 264
        Bandwidth 1000 (kbps) Burst 25000 (Bytes)
        (pkts matched/bytes matched) 0/0
        (total drops/bytes drops) 0/0

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

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-28

The **show policy-map interface** command displays the packet statistics of all classes that are configured for all service policies on the specified interface. Some of the key fields in the command output are described as follows:

Parameter	Description
`Class-map`	Class of traffic being displayed. Output is displayed for each configured class in the policy.
`offered rate`	Rate, in kbps, of packets coming in to the class.
`drop rate`	Rate, in kbps, at which packets are dropped from the class. The drop rate is calculated by subtracting the number of successfully transmitted packets from the offered rate.
`Match`	Match criteria specified for the class of traffic.
`pkts matched/bytes matched`	Number of packets (also shown in bytes) matching this class that were placed in the queue.
`depth/total drops/no-buffer drops`	Number of packets discarded for this class. No-buffer indicates that no memory buffer exists to service the packet.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **By combining basic queuing mechanisms, you can build an advanced queuing mechanism.**
- **CBWFQ is a mechanism that is used to overcome deficiencies of WFQ.**
- **CBWFQ extends the standard WFQ functionality to provide support for traffic classes. Classes are based on user-defined match criteria.**
- **CBWFQ provides a minimum bandwidth guarantee according to traffic classes.**
- **CBWFQ uses the** policy-map configuration **commands to configure such parameters as bandwidth and queue-limit within the associated class.**
- **The** show policy-map interface **is used to display the CBWFQ and LLQ statistics.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-29

Summary (Cont.)

- **LLQ extends the functionality of CBWFQ by adding a priority queue for time-sensitive traffic such as voice and video.**
- **The LLQ scheduler guarantees both low latency and bandwidth for the traffic in the priority queue.**
- **LLQ is implemented within CBWFQ by the addition of a priority queue which is serviced using a strict priority scheduler. In the event of congestion, if the priority queue traffic exceeds the bandwidth guarantee, a congestion-aware policer is used to drop the exceeding traffic.**
- **LLQ allows delay-sensitive data such as voice to be given preferential treatment over other traffic.**
- **LLQ is configured using the** priority policy-map configuration **command within the priority class.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-30

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

References

For additional information, refer to these resources:

- *Class-Based Weighted Fair Queueing*.
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt2/qcfconmg.htm#1001203
- *Low Latency Queueing*.
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt2/qcfconmg.htm#1003189
- *Configuring Weighted Fair Queueing and Low Latency Queueing*.
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt2/qcfwfq.htm
- *Configuring Class Based Weighted Fair Queueing with FRTS*.
  http://www.cisco.com/en/US/tech/tk713/tk237/technologies_configuration_example09186a008009486b.shtml

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 5

Configuring LAN Congestion Management

Overview

In a converged network, it is imperative that you implement quality of service (QoS). You need to be aware of how queuing is implemented across the Cisco Catalyst switching family, because different switches offer varied mechanisms for providing QoS. This lesson describes how several key Catalyst switches provide queuing support for QoS and how to configure those switches for QoS. The lesson also describes priority queuing (PQ), weighted round robin (WRR), shaped round robin (SRR), and weighted tail drop (WTD).

Objectives

Upon completing this lesson, you will be able to configure WRR on a Catalyst switch to manage LAN congestion. This ability includes being able to meet these objectives:

- Describe the different queuing capabilities available on Cisco Catalyst switches
- Explain how WRR works on a Catalyst 2950 switch
- Describe the commands required to configure CoS to queue mappings for PQ on Catalyst 2950 switches
- Describe the commands required to configure WRR on Catalyst 2950 switches
- Describe the commands required to monitor queuing on Catalyst 2950 switches

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Queuing on Catalyst Switches

This topic describes the different queuing capabilities available on Cisco Catalyst switches.

In a converged network, it is vital to ensure that voice traffic is not dropped. The use of multiple queues in Catalyst switches protects the queue containing important traffic (voice) from being dropped. Cisco Catalyst switches offer a variety of queuing capabilities, depending upon the switch model and port type.

One of the key options that you can assign to queues in most Catalyst switches is "drop thresholds." You can assign one or more drop thresholds to a queue. Packets are queued until the thresholds are exceeded.

For example, all packets with differentiated services code points (DSCPs) that are assigned to the first threshold are dropped until the threshold is no longer exceeded. However, packets assigned to a second threshold continue to be queued and sent as long as the second threshold is not exceeded. The thresholds are all specified as percentages ranging from 1 to 100. A value of 10 indicates a threshold in which the buffer is 10 percent full.

On some switches, queues can have normal tail drop or weighted random early detection (WRED) dropping.

Drops will occur only in data queues. The purpose of using multiple queues is to prevent voice traffic from being dropped or delayed.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Queuing on Catalyst Switches (Cont.)

- **Key queuing features depend upon the switch hardware:**
  - **The number of queues per port**
  - **The type of queues (priority or standard)**
  - **The capability to have drop thresholds for a queue**
  - **The number of drop thresholds per queue**
  - **The type of drop thresholds (tail drop or WRED)**
- **Switch queuing capabilities are shown as:**
  - **2Q2T:**
    - **Two queues**
    - **Two drop thresholds for each queue**
  - **1P2Q2T:**
    - **One priority queue**
    - **Two additional queues**
    - **Two drop thresholds for each queue**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-3

Different Cisco Catalyst switches offer these different queuing capabilities:

- The number of queues per port
- The type of queues (priority or standard)
- The capability to have drop thresholds for a queue
- The number of drop thresholds per queue
- The type of drop thresholds (tail drop or WRED).

When you view information on Cisco Catalyst switches, queuing information is displayed in an abbreviated format, as in these examples:

- **2Q2T:** Indicates that the switch supports two standard queues and two drop thresholds per queue
- **1P2Q2T:** Indicates that the switch supports one priority queue, two standard queues, and two drop thresholds per queue

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Queuing on Catalyst Switches (Cont.)

	6500	4000	3750, 2970, & 3650	3550	2950
Transmit Queues	2Q2T 1P2Q2T 1P3Q1T 1P2Q1T	1P3Q2T 4Q2T	4Q3T	1P3Q2T 4Q2T	1P3Q 4Q
Receive Queues	1Q4T 1P1Q4T 1P1Q 1P1Q8T	No	1P1Q3T 2Q3T	No	No

017G_508

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-4

The chart shows the capabilities of Catalyst switch models.

6500 Series Catalyst Switches

The Catalyst 6500 provides both receive (Rx) and transmit (Tx) queues. The number and type of queues is dependent upon the line card.

The Rx queues are designed to protect voice traffic from delays or drops. An example of the implementation of an Rx queue with a priority queue and a standard queue with drop thresholds (1P1Q4T) is as follows:

- Frames with class of service (CoS) 5 go to the priority queue.
- Frames with CoS 0, 1, 2, 3, 4, 6, or 7 go to the standard Rx queue as follows:
  - Using standard receive-queue tail-drop threshold 1, the switch drops incoming frames with CoS 0 or 1 when the receive-queue buffer is 50 percent or more full.
  - Using standard receive-queue tail-drop threshold 2, the switch drops incoming frames with CoS 2 or 3 when the receive-queue buffer is 60 percent or more full.
  - Using standard receive-queue tail-drop threshold 3, the switch drops incoming frames with CoS 4 when the receive-queue buffer is 80 percent or more full.
  - Using standard receive-queue tail-drop threshold 4, the switch drops incoming frames with CoS 6 or 7 when the receive-queue buffer is 100 percent full.

An example of the implementation of a Tx queue with two queues and two drop thresholds (2Q2T) on the Catalyst 6500 is as follows. For 2Q2T ports, each Tx queue has two tail-drop thresholds that function as follows:

- Frames with CoS 0, 1, 2, or 3 go to the low-priority Tx queue (queue 1):
  - Using Tx queue 1, tail-drop threshold 1, the switch drops frames with CoS 0 or 1 when the low-priority transmit-queue buffer is 80 percent full.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

- — Using Tx queue 1, tail-drop threshold 2, the switch drops frames with CoS 2 or 3 when the low-priority transmit-queue buffer is 100 percent full.

- Frames with CoS 4, 5, 6, or 7 go to the high-priority Tx queue (queue 2):
  - — Using Tx queue 2, tail-drop threshold 1, the switch drops frames with CoS 4 or 5 when the high-priority transmit-queue buffer is 80 percent full.
  - — Using Tx queue 2, tail-drop threshold 2, the switch drops frames with CoS 6 or 7 when the high-priority transmit-queue buffer is 100 percent full.

4500 Series Catalyst Switches

On the Catalyst 4500 with a Supervisor III engine, each physical port has four Tx queues (egress queues). Each packet that needs to be transmitted is enqueued to one of the Tx queues. The Tx queues are then serviced based on the Tx queue scheduling algorithm.

When the final transmit DSCP is computed (including any markdown of DSCP), the transmit DSCP-to-Tx queue mapping configuration determines the Tx queue. The packet is placed in the Tx queue of the transmit port, determined from the transmit DSCP.

The four Tx queues for a transmit port share the available link bandwidth of that transmit port. You can set the link bandwidth to be shared differently among the Tx queues using the **bandwidth** command in interface Tx queue configuration mode. With this command, you assign the minimum guaranteed bandwidth for each Tx queue.

By default, all queues are scheduled in a round-robin manner.

You can configure Tx queue 3 on each port with higher priority. When Tx queue 3 is configured with higher priority, packets in Tx queue 3 are scheduled ahead of packets in other queues.

When Tx queue 3 is configured at a higher priority, the packets are scheduled for transmission before the other Tx queues only if queue 3 has not met the allocated bandwidth sharing configuration—that is, if continuing to give priority to queue 3 would impact the minimum guaranteed to the other queues. Any traffic that exceeds the configured shape rate will be queued and transmitted at the configured rate. If the burst of traffic exceeds the size of the queue, packets will be dropped to maintain transmission at the configured shape rate.

Drop thresholds can be configured as tail drop or Dynamic Buffer Limiting (DBL).

3550 Catalyst Switches

On the 3550 Catalyst switches, the default scheduling method is WRR with 25 percent of bandwidth per queue by default.

You can assign CoS values to queues during configuration. The default CoS-to-queue assignment is as follows:

- CoS 6 to 7 placed in queue 4
- CoS 4 to 5 placed in queue 3
- CoS 2 to 3 placed in queue 2
- CoS 0 to 1 placed in queue 1

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The switches support WRR scheduling, and WRR with a priority queue, as follows:

- The WRR scheduling algorithm ensures that lower-priority packets are not entirely starved for bandwidth and are serviced without compromising the priority settings administered by the network manager.
- WRR with a priority queue ensures that higher-priority packets will always get serviced first, ahead of other traffic in lower-priority queues. The priority queue is defined as queue 4.

Queue weights and queue depths are configurable.

Drop thresholds can be configured as tail drop or WRED.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Queuing on Catalyst Switches (Cont.)

Catalyst 2950 Switches

- 4 transmit queues (1P3Q or 4Q)
- Need to configure PQ and ensure that CoS 5 traffic is assigned to the PQ
  - Configurable PQ for queue 4
  - Configurable CoS to specific queue
  - Configurable queue weight

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-5

On the Catalyst 2950 series switches, the default scheduling method is strict priority. Strict priority scheduling is based on the priority of queues. Packets in the high-priority queue always transmit first; packets in the low-priority queue do not transmit until all the high-priority queues become empty.

CoS values can be assigned to queues during configuration. The default CoS-to-queue assignment is:

- CoS 6 to 7 placed in queue 4
- CoS 4 to 5 placed in queue 3
- CoS 2 to 3 placed in queue 2
- CoS 0 to 1 placed in queue 1

Catalyst 2950 switches support PQ, WRR scheduling, and WRR with a priority queue, as follows:

- The WRR scheduling algorithm ensures that lower-priority packets are not entirely starved for bandwidth and are serviced without compromising the priority settings administered by the network manager.
- WRR with a priority queue ensures that higher-priority packets will always get serviced first, ahead of other traffic in lower priority queues. The priority queue is defined as queue 4.

Queue weights are configurable.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Weighted Round Robin

This topic describes how WRR works on a Catalyst 2950 switch.

Weighted Round Robin

- **WRR overcomes the problem of having PQ starving out the lower priority queues.**
- **WRR scheduling prevents queues with a lower weight from being completely starved during periods of heavy high-priority traffic.**
- **Different weights are assigned to each queue.**
- **For example, in one scheduling round, the WRR scheduler will transmit:**
  - **Three frames from a queue assigned weight 3**
  - **Four frames from a queue assigned weight 4**
- **WRR with an expedite queue: When WRR is configured on a Catalyst 2950, the option exists to configure queue 4 as a priority queue—an "expedite queue."**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-6

WRR scheduling requires that you specify a number that indicates the importance (weight) of the queue relative to the other CoS queues. WRR scheduling prevents the low-priority queues from being completely neglected during periods of high-priority traffic. The WRR scheduler transmits some packets from each queue in turn. The number of packets it sends corresponds to the relative importance of the queue. For example, if one queue has a weight of 3 and another has a weight of 4, three packets are sent from the first queue for every four that are sent from the second queue. By using this scheduling, low-priority queues have the opportunity to send packets even though the high-priority queues are not empty.

WRR scheduling with an expedite priority queue (also referred to as strict PQ) uses one of the egress queues as an expedite queue (queue 4 on a Catalyst 2950). The remaining queues participate in WRR. When the expedite queue is configured, it is a priority queue and is serviced until it is empty before the other queues are serviced by WRR scheduling. Actions at the egress interface include queuing and scheduling, as follows:

- Queuing evaluates the internal DSCP and determines in which of the four egress queues to place the packet. The DSCP value is mapped to a CoS value, which selects one of the queues.
- Scheduling services the four egress queues based on their configured WRR weights and thresholds. One of the queues can be the expedite queue, which is serviced until empty before the other queues are serviced.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

With WRR, lower-priority queues have the opportunity to transmit packets even though the high-priority queues have not been emptied. With WRR with an expedite queue, one queue (queue 4 on the Catalyst 2950 and 3550) can be configured as an expedite priority queue. All traffic from the expedite queue must be serviced before the remaining three queues are serviced. The expedite queue can be used to ensure that voice traffic incurs minimal delay and no drops.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring CoS-to-Queue Mappings for PQ on Catalyst 2950 Switches

This topic describes the commands required to configure CoS-to-queue mappings for PQ on the Cisco Catalyst 2950 switch.

Configuring PQ on Catalyst 2950 Switches

```
Switch(config)#
wrr-queue cos-map quid cos1...cosn
```

- **Assigns CoS values to CoS priority queues**
- **quid: Specifies the queue ID of the CoS priority queue. (Ranges are 1 to 4 where 1 is the lowest CoS priority queue.)**
- **cos1...cos*n*: Specifies the CoS values that are mapped to the queue ID.**
- **Default ID values are:**

Queue ID	CoS Values
1	0, 1
2	2, 3
3	4, 5
4	6, 7

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-7

To configure CoS-to-queue mappings for PQ on the Catalyst 2950 switch, specify the queue ID of the CoS priority queue. Ranges are 1 to 4 where 1 is the lowest CoS priority queue. Then, specify the CoS values that are mapped to the queue ID.

wrr-queue cos-map *quid cos1...cosn*

Syntax Description

Parameter	Description
quid	The queue ID of the CoS priority queue. Ranges are 1 to 4 where 1 is the lowest CoS priority queue.
cos1...cosn	The CoS values that are mapped to the queue ID.

The default CoS-to-priority queue assignments are shown in this table.

Queue	1	2	3	4
CoS Values	0, 1	2, 3	4, 5	6, 7

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring WRR on Catalyst 2950 Switches

This topic describes the commands required to configure WRR on the Catalyst 2950 switch.

Configuring WRR on Catalyst 2950 Switches

```
Switch(config-if)#
wrr-queue bandwidth weight1...weight4
```

- **Assigns WRR weights to the four egress queues**
- **Ranges for the WRR values:**
  - **For `weight1`, `weight2`, and `weight3`, the range is 1 to 255.**
  - **For `weight4`, the range is 0 to 255 (when `weight4` is set to 0, queue 4 is configured as the expedite queue).**

```
!
interface GigabitEthernet0/12
  wrr-queue bandwidth 20 1 80 0
  no wrr-queue cos-map
  wrr-queue cos-map 1 0 1 2 4
  wrr-queue cos-map 3 3 6 7
  wrr-queue cos-map 4 5
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-8

Use the **wrr-queue bandwidth** global configuration command to assign WRR weights to the four CoS priority queues on the Catalyst 2950 switch. Use the **no** form of this command to disable the WRR scheduler and enable the strict priority scheduler.

For weight 1, weight 2, and weight 3, the range is 1 to 255. The range for weight 4 is 0 to 255. Queues 1, 2, and 3 can be configured for WRR scheduling and queue 4 can be configured for strict priority scheduling. To configure queue 4 as the expedite queue, set weight 4 to 0. When queue 4 is empty, packets from queues 1, 2, and 3 are sent according to the assigned WRR weights.

wrr-queue bandwidth *weight1...weight4*

Syntax Description

Parameter	Description
`weight1...weight4`	The ratio of weight 1, weight 2, weight 3, and weight 4 determines the weights of the WRR scheduler.

Note In Cisco IOS software releases earlier than Release 12.1(12c)EA1, the range for all queues is 1 to 255.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

In the example for the Catalyst 2950 switch, this configuration has been made:

- The interface is set to GigabitEthernet 0/12.
- Queue bandwidth (queue weight) is set to these weights:
  - Queue 1: 20
  - Queue 2: 1
  - Queue 3: 80
  - Queue 4: 0 (because this is 0, this is the expedite queue)
- The CoS map is set to its default settings.
- CoS is mapped to queues according to the following:
  - Queue 1: CoS 0, 1, 2, 4
  - Queue 2: No CoS assigned
  - Queue 3: CoS 3, 6, 7
  - Queue 4: CoS 5 (voice traffic goes to the expedite queue)

Note This is the AutoQoS configuration for the Catalyst 2950.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring Queuing on Catalyst 2950 Switches

This topic describes the commands required to monitor queuing on Catalyst 2950 switches.

Monitoring Queuing on Catalyst 2950 Switches

```
Switch>
show mls qos maps [cos-dscp | dscp-cos]
```

- **Displays QoS mapping information.**
- **This command is available with enhanced software image switches.**

```
Switch> show mls qos maps

Dscp-cos map:
dscp: 0 8 10 16 18 24 26 32 34 40 42 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 40 48 56
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-9

Use the **show mls qos maps** command to display QoS mapping information on the Catalyst 2950 switch. Maps are used to generate an internal DSCP value, which represents the priority of the traffic.

The **show mls qos maps** command is available only if the Catalyst 2950 switch is running the Enhanced Image (EI) software.

If the **show mls qos maps** command is used without any keywords, it will display all maps.

show mls qos maps [cos-dscp | dscp-cos]

Syntax Description

Parameter	Description
`cos-dscp`	(Optional) Displays CoS-to-DSCP map.
`dscp-cos`	(Optional) Displays DSCP-to-CoS map.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring Queuing on Catalyst 2950 Switches (Cont.)

```
Switch>
show wrr-queue bandwidth
```

- **Displays the WRR bandwidth allocation for the CoS priority queues**

```
Switch> show wrr-queue bandwidth
WRR Queue : 1 2 3 4
Bandwidth : 10 20 30 40
```

```
Switch>
show wrr-queue cos-map
```

- **Displays the mapping of the CoS priority queues**

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

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-10

The **show wrr-queue bandwidth** command is used to display the WRR bandwidth allocation for the four CoS priority queues.

The **show wrr-queue cos-map** command is used to display the mapping of the CoS priority queues.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring Queuing on Catalyst 2950 Switches (Cont.)

```
Switch>
show mls qos interface [interface-id] [policers]
```

- **Displays QoS information at the interface level**

```
Switch> show mls qos interface fastethernet0/1
FastEthernet0/1
trust state:trust cos
trust mode:trust cos
COS override:dis
default COS:0
pass-through:none
trust device:cisco-phone
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-11

Use the **show mls qos interface** command to display QoS information at the interface level. Although it will be visible in command-line interface (CLI) help strings, the **policers** keyword is available only when the Catalyst 2950 switch is running the enhanced software image.

show mls qos interface [*interface-id***] [policers]**

Syntax Description

Parameter	Description
interface-id	(Optional) Displays QoS information for the specified interface.
policers	(Optional) Displays all the policers configured on the interface, their settings, and the number of policers unassigned. Available only when the switch is running the EI software.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **The number and capabilities of queues on Catalyst switches depend upon the model of the switch, supervisor, and line cards. PQ and WRR are the two queuing methods used for Catalyst switches. The use of PQ can starve lower-priority queues.**
- **With WRR, different weights are assigned to each queue. Use of WRR scheduling prevents the low-priority queues from being completely neglected during periods of high-priority traffic. On most Catalyst switches, a single priority queue can be configured with WRR to ensure priority dispatch of voice traffic.**
- **To configure CoS-to-queue mappings for PQ on the Catalyst 2950 switch, specify the queue ID of the CoS priority queue. (Ranges are 1 to 4 where 1 is the lowest CoS priority queue.) Then, specify the CoS values that are mapped to the queue ID. Use the** wrr-queue cos-map *quid cos1...cosn* **command.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-12

Summary (Cont.)

- **The** wrr-queue bandwidth **global configuration command is used to assign WRR weights to the four CoS priority queues on the Catalyst 2950 switch.**
- **The** show mls qos maps **command is used to display QoS mapping information on the Catalyst 2950 switch.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-13

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Summary

This topic summarizes the key points discussed in this module.

Module Summary

- **Congestion can occur at any point in the network, but particularly at points of speed mismatches and traffic aggregation. Queuing algorithms such as FIFO, priority, and round robin are used to manage congestion.**
- **Each physical interface has a hardware and a software queuing system.**
- **WFQ was developed to overcome the limitations of the more basic queuing methods.**
- **CBWFQ extends the standard WFQ functionality to provide support for user-defined traffic classes. LLQ extends the functionality of CBWFQ by adding priority queues for time-sensitive traffic such as voice and video.**
- **PQ, WRR, and WRR with a PQ are the three key queuing methods used for Catalyst switches.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—5-1

Effective congestion management is key to QoS in converged networks. Low-latency traffic such as voice and video must be constantly moved to high-priority queues in order to ensure reasonable quality.

Cisco routers offer a variety of simple (FIFO, PQ, and CQ) and sophisticated (WFQ, CBWFQ, and LLQ) queuing algorithms to provide effective congestion management on converged networks. LLQ, the most sophisticated queuing algorithm, was specifically designed to provide the highest QoS to voice traffic.

Cisco switches offer a variety of queuing capabilities, depending upon the model of switch being used. On most Catalyst switches, three queuing methods are available for use: PQ, WRR, and WRR with a priority queue.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

References

For additional information, refer to these resources:

- To learn more about congestion and queuing, refer to *Understanding Delay in Packet Voice Networks* at this URL: http://www.cisco.com/en/US/tech/tk652/tk698/technologies_white_paper09186a00800a8993.shtml
- To learn more about congestion and queuing, refer to *Understanding Jitter in Packet Voice Networks (Cisco IOS Platforms)* at this URL: http://www.cisco.com/en/US/tech/tk652/tk698/technologies_tech_note09186a00800945df.shtml
- To learn more about configuring WFQ, refer to *Configuring Weighted Fair Queueing* at this URL: http://www.cisco.com/en/US/products/sw/iosswrel/ps1828/products_configuration_guide_chapter09186a00800ca597.html
- To learn more about queuing on the Cisco Catalyst 3550 series switches, refer to *Configuring QoS* at this URL: http://www.cisco.com/en/US/products/hw/switches/ps646/products_configuration_guide_chapter09186a008014f36e.html#1127419
- To learn more about queuing on the Cisco Catalyst 2950 series switches, refer to *Configuring QoS* at this URL: http://www.cisco.com/en/US/products/hw/switches/ps628/products_configuration_guide_chapter09186a008014f2c0.html#1025310
- To learn more about queuing on the Cisco Catalyst 4000 series switches, refer to *Configuring QoS* at this URL: http://www.cisco.com/en/US/products/hw/switches/ps663/products_configuration_guide_chapter09186a008007eddd.html
- To learn more about queuing on the Cisco Catalyst 6500 series switches, refer to *Configuring QoS* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/122sx/swcg/qos.htm

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check

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

Q1) What happens when the highest-priority queue becomes congested in a priority queuing algorithm? (Source: Introducing Queuing)

A) All the other queues starve.
B) Tail dropping focuses on the highest-priority queue.
C) Other queues are served on a round-robin basis.
D) Packets in the highest-priority queue are moved to a lower-priority queue.

Q2) In WRR implementation, using a byte threshold as a measurement of each queue share of bandwidth, given an MTU of 2000 and a byte-count of 4000, what would the router do with the next packet for queue 2 (800 bytes) if the router had just dispatched two packets from queue 2 (sizes 2000 and 1600) to the hardware queue? (Source: Introducing Queuing)

A) tail drop the next packet
B) dispatch the first packet from the next queue
C) dispatch the next packet to the hardware queue
D) split the packet and transmit the first 400 bytes

Q3) On a network device connecting a LAN and a WAN, what type of congestion would be likely for traffic moving from the LAN to the WAN? (Source: Understanding Queuing Implementations)

A) bursty
B) transient
C) persistent
D) fluctuating

Q4) Given that the hardware queue is *not* full, how will the next packet be serviced by the software queue? (Source: Understanding Queuing Implementations)

A) software queue will be bypassed
B) software queue will enqueue the packet
C) software queue will expedite the packet
D) software queue will only meter the packet

Q5) How does WFQ implement tail dropping? (Source: Configuring FIFO and WFQ)

A) drops the last packet to arrive
B) drops all nonvoice packets first
C) drops the lowest-priority packets first
D) drops packets from the most aggressive flows

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q6) Consider that a WFQ system has a modest hold-queue limit of ten (HQO = 10) and a congestive discard threshold of eight (CDT = 8), and that there are already eight packets in the system. If a newly arriving packet had the worst finish time of all packets in the system, what would happen to the packet? (Source: Configuring FIFO and WFQ)

A) it would be dropped
B) it would be enqueued
C) it would be buffered until a spot in a queue came open
D) it would be dispatched

Q7) Which of the following is the default dropping scheme for CBWFQ? (Source: Configuring CBWFQ and LLQ)

A) RED
B) WRED
C) tail drop
D) class-based policing

Q8) What does LLQ bring to CBWFQ? (Source: Configuring CBWFQ and LLQ)

A) strict priority scheduling
B) alternate priority scheduling
C) non-policed queues for low-latency traffic
D) special voice traffic classification and dispatch

Q9) Which type of traffic should you limit the use of the "priority" command to? (Source: Configuring CBWFQ and LLQ)

A) low-latency data traffic
B) voice-like CBR traffic
C) high-volume VBR traffic
D) video and teleconferencing, available (ABR) traffic

Q10) When WRR with an expedite queue has been configured on a Cisco Catalyst 2950 switch, which queue is emptied before any other queues are serviced? (Source: Configuring LAN Congestion Management)

A) queue 1
B) queue 2
C) queue 3
D) queue 4

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check Answer Key

Q1) A

Q2) C

Q3) C

Q4) A

Q5) D

Q6) A

Q7) C

Q8) A

Q9) B

Q10) D

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module 6

Congestion Avoidance

Overview

Congestion is a normal occurrence in networks. Whether congestion occurs as a result of a lack of buffer space, network aggregation points, or a low-speed wide-area link, many congestion management techniques exist to ensure that specific applications and traffic classes are given their share of available bandwidth when congestion occurs. Congestion management does, however, come at a price. When congestion occurs, some traffic is delayed or even dropped at the expense of other traffic. When drops occur, different problems may arise which can exacerbate the congestion, such as retransmissions and TCP global synchronization in TCP/IP networks. Network administrators can use congestion avoidance mechanisms to reduce the negative effects of congestion by penalizing the most aggressive traffic streams as software queues begin to fill. This module describes the problems with TCP congestion management and the benefits of deploying congestion avoidance mechanisms in a network.

Module Objectives

Upon completing this module, you will be able to use Cisco quality of service (QoS) congestion avoidance mechanisms to reduce the effects of congestion on the network. This ability includes being able to meet these objectives:

- Explain the problems that may result from the limitations of TCP congestion management mechanisms on a converged network
- Explain how RED can be used to avoid congestion
- Configure CB-WRED to avoid congestion
- Configure ECN to enhance the congestion avoidance features of WRED

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 1

Introducing Congestion Avoidance

Overview

TCP supports traffic management mechanisms such as slow start and fast retransmit. When congestion occurs, tail-dropping the TCP traffic can cause TCP global synchronization, resulting in poor bandwidth use. This lesson describes how TCP manages the traffic flow between two hosts, and the effects of tail-dropping on TCP traffic.

Objectives

Upon completing this lesson, you will be able to explain the problems that may result from the limitations of TCP congestion management mechanisms on a converged network. This ability includes being able to meet these objectives:

- Explain the behavior of TCP senders and receivers when sending packets
- Explain how TCP responds to congestion
- Describe the default mechanism for managing interface congestion with tail drop
- Describe the limitations of using tail drop as a congestion management mechanism

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Behavior of TCP Senders and Receivers

This topic describes the behavior of TCP senders and receivers when sending packets.

Before any data is transmitted using TCP, a connection must first be established between the transmitting and receiving hosts. When the connection is initially established, the two hosts must agree on certain parameters that will be used during the communication session. One of the parameters that must be decided is called the window size, or how many data bytes to transmit at a time. Initially, TCP sends a small number of data bytes, and then exponentially increases the number sent. For example, a TCP session originating from host A begins with a window size of 1 and therefore sends one packet. When host A receives a positive ACK from the receiver, Host A increases its window size to 2. Host A then sends 2 packets, receives a positive ACK, and increases its window size to 4, and so on.

Note TCP tracks window size by byte count. For purposes of illustration packets, N is used.

In traditional TCP, the maximum window size is 64 KB (65,535 bytes). Extensions to TCP, specified in RFC 1323, allow for tuning TCP by extending the maximum TCP window size to 2^{30} bytes. TCP extensions for high performance, although supported on most operating systems, may not be supported on your system.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Windowing in TCP

After connecting to a website, a file transfer using the FTP download is initiated. Watching the progress of the transfer, the bytes-per-second counter steadily increases during the file transfer. This is an example of TCP windowing in action.

When the receiver receives a data segment, the receiver checks that data segment sequence number (byte count). If the data received fills in the next sequence of numbers expected, the receiver indicates that the data segment was received in order. The receiver then delivers all the data that it holds to the target application, and updates the sequence number to reflect the next byte number in expected order.

When this process is complete, the receiver performs one of these actions:

- Immediately transmits an acknowledgment (ACK) to the sender
- Schedules an ACK to be transmitted to the sender after a short delay

The ACK notifies the sender that the receiver received all data segments up to but not including the byte number in the new sequence number. Receivers usually try to send an ACK in response to alternating data segments they receive. They send the ACK because, for many applications, if the receiver waits out a small delay, it can efficiently piggyback its reply acknowledgment on a normal response to the sender. However, when the receiver receives a data segment out of order, it immediately responds with an ACK to direct the sender to retransmit the lost data segment.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Congestion and TCP

This topic describes the TCP response to congestion.

When the sender receives an ACK, the sender determines if any data is outstanding:

- If no data is outstanding, the sender determines that the ACK is a keepalive, meant to keep the line active, and it does nothing.
- If data is outstanding, the sender determines whether the ACK indicates that the receiver has received some or none of the data.
  - If the ACK acknowledges receipt of some data sent, the sender determines if new credit has been granted to allow it to send more data.
  - When the ACK acknowledges receipt of none of the sent data and there is outstanding data, the sender interprets the ACK to be a repeatedly sent ACK. This condition indicates that some data was received out of order, forcing the receiver to remit the first ACK, and that a second data segment was received out of order, forcing the receiver to remit the second ACK. In most cases, the receiver would receive two segments out of order, because one of the data segments had been dropped.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

When a TCP sender detects a dropped data segment, it retransmits the segment. Then the sender slows its transmission rate so that the rate is half of what it was before the drop was detected. This is known as the TCP slow-start mechanism. In the figure, a station transmits three packets to the receiving station. Unfortunately, the first packet is dropped somewhere in the network. Therefore the receiver sends an ACK 1, to request the missing packet. Because the transmitter does not know if the ACK was just a duplicate ACK, it will wait for three ACK 1 packets from the receiver. Upon receipt of the third ACK, the missing packet, packet 1, is resent to the receiver. The receiver now sends an ACK 4 indicating that it has already received packets 2 and 3 and is ready for the next packet.

Although the TCP slow-start behavior is appropriately responsive to congestion, problems can arise when multiple TCP sessions are concurrently carried on the same router and all TCP senders slow down transmission of packets at the same time.

If a TCP sender does not receive acknowledgement for sent segments, it cannot wait indefinitely before it assumes that the data segment that was sent never arrived at the receiver. TCP senders maintain the retransmission timer to trigger a segment retransmission. The retransmission timer can impact TCP performance. If the retransmission timer is too short, duplicate data will be sent into the network unnecessarily. If the retransmission timer is too long, the sender will wait (remain idle) for too long, slowing down the flow of data.

The selective acknowledgment (SACK) mechanism, as proposed in RFC 2018, can improve the time it takes for the sender to recover from multiple packet losses, because noncontiguous blocks of data can be acknowledged, and the sender only has to retransmit data that is actually lost. SACK is used to convey extended acknowledgement information from the receiver to the sender to inform the sender of noncontiguous blocks of data that have been received. Using the example in the slide, instead of sending back an ACK N + 1, the receiver can send a SACK N + 1 and also indicate back to the sender that N + 3 has been correctly received with the SACK option.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

In standard TCP implementations, a TCP sender can only discover that a single packet has been lost each round-trip time (RTT), causing poor TCP performance when multiple packets are lost.

Remember, the sender must receive three duplicate ACK packets before it realizes that a packet has been lost. As a result of receiving the third ACK, the sender will immediately send the segment referred to by the ACK. This TCP behavior is called fast retransmit.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Managing Interface Congestion with Tail Drop

This topic describes the default mechanism for managing interface congestion with tail drop.

When an interface on a router cannot transmit a packet immediately, the packet is queued, either in an interface transmit (Tx) ring, or the interface output hold queue, depending on the switching path that is used. Packets are then taken out of the queue and eventually transmitted on the interface.

If the arrival rate of packets to the output interface exceeds the router capability to buffer and forward traffic, the queues increase to their maximum length and the interface becomes congested. Tail drop is the default queuing response to congestion. Tail drop treats all traffic equally and does not differentiate between classes of service. Applications may suffer performance degradation due to packet loss caused by tail drop. When the output queue is full and tail drop is in effect, all packets trying to enter (at the tail of) the queue are dropped until the congestion is eliminated and the queue is no longer full.

Weighted fair queuing (WFQ), if configured on an interface, has a more elaborate scheme for dropping traffic. It is able to punish the most aggressive flows via its congestive discard threshold (CDT)-based dropping algorithm. Unfortunately, WFQ does not scale to backbone speeds.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Tail Drop Limitations

This topic describes the limitations of using tail drop as a congestion management mechanism.

Tail Drop Limitations

- **Tail drop should be avoided because it contains significant flaws:**
  - **TCP synchronization**
  - **TCP starvation**
  - **No differentiated drop**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-7

The simple tail-drop scheme does not work well in environments with a large number of TCP flows or in environments in which selective dropping is required. Understanding the network interaction between TCP stack intelligence and dropping is required to implement a more efficient and fair dropping scheme, especially in service provider environments.

Tail drop has the following shortcomings:

- When congestion occurs, dropping affects most of the TCP sessions, which simultaneously back off and then restart again. This causes inefficient link utilization at the congestion point (TCP global synchronization).
- TCP starvation, in which all buffers are temporarily seized by aggressive flows, and normal TCP flows experience buffer starvation.
- There is no differentiated drop mechanism, and therefore premium traffic is dropped in the same way as best-effort traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

TCP Synchronization

- **Multiple TCP sessions start at different times.**
- **TCP window sizes are increased.**
- **Tail drops cause many packets of many sessions to be dropped at the same time.**
- **TCP sessions restart at the same time (synchronized).**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-8

A router can handle multiple concurrent TCP sessions. It is likely that when traffic exceeds the queue limit, it exceeds this limit due to the bursty nature of packet networks. However, there is also a high probability that excessive traffic depth caused by packet bursts are temporary and that traffic does not stay excessively deep except at points where traffic flows merge, or at edge routers.

If the receiving router drops all traffic that exceeds the queue limit, as is done with tail drop by default, many TCP sessions simultaneously go into slow start. Consequently, traffic temporarily slows down to the extreme and then all flows slow-start again. This activity creates a condition called global synchronization.

Global synchronization occurs as waves of congestion crest only to be followed by troughs during which the transmission link is not fully used. Global synchronization of TCP hosts can occur because packets are dropped all at once. Global synchronization occurs when multiple TCP hosts reduce their transmission rates in response to packet dropping. When congestion is reduced, their transmission rates are increased. The most important point is that the waves of transmission known as global synchronization result in significant link underutilization.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

TCP Delay, Jitter and Starvation

- **Constant high buffer usage (long queue) causes delay.**
- **More aggressive flows can cause other flows to starve.**
- **No differentiated dropping occurs.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-9

During periods of congestion, packets are queued up to the full queue length, which also causes increased delay for packets that are already in the queue. In addition, queuing introduces unequal delays for packets of the same flow, thus producing jitter.

Another TCP-related phenomenon that reduces optimal throughput of network applications is TCP starvation. When multiple flows are established over a router, some of these flows may be much more aggressive than other flows. For instance, when a file transfer application TCP transmit window increases, the TCP session can send a number of large packets to its destination. The packets immediately fill the queue on the router, and other, less aggressive flows can be starved because there is no differentiated treatment indicating which packets should be dropped. As a result, these less aggressive flows are tail-dropped at the output interface.

Based on the knowledge of TCP behavior during periods of congestion, you can conclude that tail drop is not the optimal mechanism for congestion avoidance and therefore should not be used. Instead, more intelligent congestion avoidance mechanisms should be used that slow down traffic before actual congestion occurs.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **TCP sender sends as many bytes as credit allows. TCP receiver acknowledges the next segment it expects to receive.**
- **TCP uses windowing and the TCP slow-start mechanism as its means of controlling congestion.**
- **By default, routers resort to tail drop, relying on TCP congestion controls when queues become full.**
- **Tail drop causes significant issues including TCP synchronization, starvation, and delay. TCP synchronization decreases the average utilization of network links.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-10

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 2

Introducing RED

Overview

Congestion avoidance techniques monitor network traffic loads in an effort to anticipate and avoid congestion at common network bottleneck points. Congestion avoidance is achieved through packet dropping using a more complex dropping technique than simple tail drop.

This lesson describes the congestion avoidance technique random early detection (RED) and its scalable dropping method, which is suitable for low- and high-speed networks.

Objectives

Upon completing this lesson, you will be able to explain how RED can be used to avoid congestion. This ability includes being able to meet these objectives:

- Describe RED and how it can be used to prevent congestion
- Describe the elements of a RED traffic profile used to implement the RED packet-dropping strategy
- Describe the different packet drop modes of RED
- Describe the effects of using RED on TCP traffic by comparing TCP traffic flows both before and after the application of RED
- Identify the points in a network where congestion avoidance mechanisms are commonly deployed in enterprise and service provider networks

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Random Early Detection

This topic describes RED and how it can be used to prevent congestion.

Random Early Detection

- **Tail drop can be avoided if congestion is prevented.**
- **RED is a mechanism that randomly drops packets before a queue is full.**
- **RED increases drop rate as the average queue size increases.**
- **RED result:**
  - **TCP sessions slow down to the approximate rate of output-link bandwidth.**
  - **Average queue size is small (much less than the maximum queue size).**
  - **TCP sessions are desynchronized by random drops.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-2

RED is a dropping mechanism that randomly drops packets before a queue is full. The dropping strategy is based primarily on the average queue length—that is, when the average size of the queue increases, RED will be more likely to drop an incoming packet than when the average queue length is shorter.

Because RED drops packets randomly, it has no per-flow intelligence. The rationale is that an aggressive flow will represent most of the arriving traffic and therefore it is likely that RED will drop a packet of an aggressive session. RED therefore punishes more aggressive sessions with higher statistical probability and is, therefore, able to somewhat selectively slow down the most significant cause of congestion. Directing one TCP session at a time to slow down allows for full utilization of the bandwidth, rather than utilization that manifests itself as crests and troughs of traffic.

As a result of implementing RED, the problem of TCP global synchronization is much less likely to occur, and TCP can utilize link bandwidth more efficiently. In RED implementations, the average queue size also decreases significantly, as the possibility of the queue filling up is reduced. This is because of very aggressive dropping in the event of traffic bursts, when the queue is already quite full.

RED distributes losses over time and normally maintains a low queue depth while absorbing traffic spikes. RED can also utilize IP precedence or differentiated services code points (DSCPs) to establish different drop profiles for different classes of traffic, in which case it is referred to as weighted random early detect (WRED) or DSCP-based WRED, respectively.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

RED Profiles

This topic describes the elements of a RED traffic profile used to implement the RED packet-dropping strategy.

A RED traffic profile is used to determine the packet-dropping strategy and is based on the average queue length. The probability of a packet being dropped is based on three configurable parameters contained within the RED profile:

- **Minimum threshold:** When the average queue length is equal or above the minimum threshold, RED starts dropping packets. The rate of packet drop increases linearly as the average queue size increases, until the average queue size reaches the maximum threshold.
- **Maximum threshold:** When the average queue size is above the maximum threshold, all packets are dropped.
- **Mark probability denominator:** This is the fraction of packets that are dropped when the average queue depth is at the maximum threshold. For example, if the denominator is 512, one out of every 512 packets is dropped when the average queue is at the maximum threshold. The linear increase of packet drops from the minimum threshold (0 drops) to the maximum threshold is based on this parameter and the queue size between the minimum and maximum thresholds.

The minimum threshold value should be set high enough to maximize the link utilization. If the minimum threshold is too low, packets may be dropped unnecessarily, and the transmission link will not be fully used.

The difference between the maximum threshold and the minimum threshold should be large enough to avoid global synchronization. If the difference is too small, many packets may be dropped at once, resulting in global synchronization.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The mark probability has the effect of controlling the number of packets that are dropped when the average queue length reaches the maximum threshold. If the value is set too low, it will result in too many dropped packets. If the value is set too large, RED dropping can be rendered ineffective.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

RED Modes

This topic describes the different packet-drop modes of RED.

RED Modes

- **RED has three modes:**
  - **No drop: When the average queue size is between 0 and the minimum threshold**
  - **Random drop: When the average queue size is between the minimum and the maximum threshold**
  - **Full drop (tail drop): When the average queue size is at maximum threshold or above**
- **Random drop should prevent congestion (prevent tail drops)**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-4

Based on the average queue size, RED has three dropping modes:

- When the average queue size is between 0 and the configured minimum threshold, no drops occur and all packets are queued.
- When the average queue size is between the configured minimum threshold and the configured maximum threshold, random drops occur, which is linearly proportional to the mark probability denominator and the average queue length.
- When the average queue size is at or higher than the maximum threshold, RED performs full (tail) drop in the queue. This is unlikely, as RED should slow down TCP traffic ahead of congestion. If a lot of non-TCP traffic is present, RED cannot effectively drop traffic to reduce congestion, and tail drops are likely to occur.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

TCP Traffic Before and After RED

This topic describes the effects of using RED on TCP traffic by comparing TCP traffic flows both before and after the application of RED.

TCP Traffic Before RED

- **TCP synchronization prevents average link utilization close to the link bandwidth.**
- **Tail drops cause TCP sessions to go into slow-start.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-5

The figure shows TCP throughput behavior compared to link bandwidth in a congested network scenario where the tail-drop mechanism is in use on a router. The global synchronization phenomenon causes all sessions to slow down when congestion occurs. All sessions are penalized when tail drop is used because it drops packets with no discrimination between individual flows.

When all sessions slow down, congestion on the router interface is removed and all TCP sessions restart their transmission at roughly the same time. Again, the router interface quickly becomes congested, causing tail drop. As a result, all TCP sessions back off again. This behavior cycles constantly, resulting in a link that is generally underutilized.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

TCP Traffic After RED

- **Average link utilization is much closer to link bandwidth.**
- **Random drops cause TCP sessions to reduce window sizes.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-6

This figure shows TCP throughput behavior compared to link bandwidth in a congested network scenario in which RED has been configured on a router. RED randomly drops packets, influencing a small number of sessions at a time, before the interface reaches congestion. Overall throughput of sessions is increased, as well as average link utilization. Global synchronization is very unlikely to occur, due to selective, but random, dropping of adaptive traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Applying Congestion Avoidance

This topic describes the points in a network where congestion avoidance mechanisms are commonly deployed in enterprise and service provider networks.

RED is most useful in enterprise and service provider networks on output interfaces where congestion is expected to occur. Edge routers or switches typically classify and mark packets as the packets enter the network. Congestion avoidance mechanisms can use these packet markings to indicate a set of drop criteria for a traffic stream.

Congestion avoidance mechanisms are also applicable to the campus or LAN environment. In these networks, congestion avoidance is best used on interfaces that connect to WAN gateways, because these interfaces are typically sites for congestion to occur.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **RED is a mechanism that randomly drops packets before a queue is full, preventing congestion and avoiding tail drop.**
- **RED operates by increasing the rate at which packets are dropped from queues as the average queue size increases.**
- **RED has three modes of operation: no drop, random drop, and full drop (tail drop).**
- **With RED, TCP global synchronization is eliminated and the average link utilization increases.**
- **RED is most useful in enterprise and service provider networks on output interfaces where congestion is expected to occur. Edge routers or switches typically classify and mark packets as they enter the network.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-8

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 3

Configuring Class-Based Weighted RED

Overview

Congestion avoidance techniques monitor network traffic loads in an effort to anticipate and avoid congestion at common network bottleneck points using advanced packet discard techniques.

This lesson describes the weighted random early detection (WRED) congestion avoidance technique, which is the Cisco implementation of random early detection (RED).

Objectives

Upon completing this lesson, you will be able to configure class-based weighted random early detection (CB-WRED) to avoid congestion. This ability includes being able to meet these objectives:

- Describe WRED and how it can be used to prevent congestion
- Describe the traffic profiles used in WRED implementations
- Identify the Cisco IOS commands required to configure CB-WRED
- Identify the Cisco IOS commands required to configure DSCP-based CB-WRED
- Identify the Cisco IOS commands used to monitor CB-WRED

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Weighted Random Early Detection

This topic describes WRED and how you can use it to prevent congestion.

Weighted Random Early Detection

- WRED can use multiple different RED profiles.
- Each profile is identified by:
  - Minimum threshold
  - Maximum threshold
  - Maximum drop probability
- WRED profile selection is based on:
  - IP precedence (8 profiles)
  - DSCP (64 profiles)
- WRED drops less important packets more aggressively than more important packets.
- WRED can be applied at the interface, VC, or class level.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-2

WRED combines RED with IP precedence or differentiated services code point (DSCP) and performs packet dropping based on IP precedence (WRED) or DSCP (DSCP-based WRED).

As with RED, WRED monitors the average queue length in the router and determines when to begin discarding packets based on the length of the interface queue. When the average queue length is greater than the user-specified minimum threshold, WRED begins to randomly drop packets (both TCP and User Datagram Protocol [UDP] packets) with a certain probability. If the average length of the queue continues to increase so that it becomes larger than the user-specified maximum threshold, WRED reverts to a tail-drop packet discard strategy, in which all incoming packets may be dropped.

The idea behind using WRED is to maintain the queue length at a level somewhere between the minimum and maximum thresholds, and to implement different drop policies for different classes of traffic. WRED can selectively discard lower-priority traffic when the interface becomes congested, and can provide differentiated performance characteristics for different classes of service. You can also configure WRED to achieve nonweighted RED behavior.

For interfaces configured to use the Resource Reservation Protocol (RSVP), WRED chooses packets from other flows to drop rather than the RSVP flows. Also, IP precedence or DSCP governs which packets are dropped, because traffic at a lower priority has a higher drop rate than traffic at a higher priority (and, therefore, lower priority is more likely to be slowed). In addition, WRED statistically drops more packets from large users than small users. Therefore, traffic sources that generate the most traffic are more likely to be slowed down than traffic sources that generate little traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WRED reduces the chances of tail drop by selectively dropping packets when the output interface begins to show signs of congestion. By dropping some packets early rather than waiting until the queue is full, WRED avoids dropping large numbers of packets at once and minimizes the chances of global synchronization. As a result, WRED maximizes the utilization of transmission lines.

WRED is only useful when the bulk of the traffic is TCP traffic. With TCP, dropped packets indicate congestion, so the packet source reduces its transmission rate. With other protocols, packet sources might not respond or might re-send dropped packets at the same rate, and so dropping packets might not decrease congestion.

WRED treats non-IP traffic as precedence 0, the lowest precedence. Therefore, non-IP traffic, in general, is more likely to be dropped than IP traffic.

WRED should be used wherever there is a potential bottleneck, or congested link—this could possibly be an access or edge link. However, WRED is normally used in the core routers of a network rather than at the network edge. Edge routers assign IP precedence or DSCP to packets as they enter the network. WRED uses these assigned values to determine how to treat different types of traffic.

WRED is not recommended for any voice queue, although WRED may be enabled on an interface carrying voice traffic. WRED will not throttle back voice traffic because voice traffic is UDP-based. The network itself should be designed not to lose voice packets because lost voice packets result in reduced voice quality. WRED controls congestion by impacting prioritized traffic other than voice, and avoiding congestion helps to ensure voice quality.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based WRED

- **Class-based WRED is available when configured in combination with CBWFQ only.**
- **Using CBWFQ with WRED allows the implementation of DiffServ Assured Forwarding PHB.**
- **Class-based configuration of WRED is identical to standalone WRED.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-3

Congestion avoidance techniques monitor the network interface load in an effort to anticipate and avoid congestion at common network bottlenecks. Congestion avoidance is achieved through intelligent packet dropping techniques. Traditionally, Cisco IOS software used standalone RED and WRED mechanisms to avoid congestion on an interface. Those mechanisms can perform a differentiated drop based on the IP precedence or DSCP value.

The class-based weighted fair queuing (CBWFQ) system supports the use of WRED inside the queuing system, thereby implementing CB-WRED. Each class is queued in its separate queue and has a queue limit, performing tail drop by default. WRED can be configured as the preferred dropping method in a queue, implementing a differentiated drop based on traffic class, and further, on the IP precedence or DSCP value.

Note The combination of CBWFQ and WRED on a single device is currently the only way to implement the DiffServ Assured Forwarding per-hop behavior (AF PHB) using Cisco IOS software.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The figure shows how WRED is implemented, and the parameters that are used by WRED to influence packet-drop decisions.

The router constantly updates the WRED algorithm with the calculated average queue length, which is based on the recent history of queue lengths.

Configured in the traffic profile are the parameters that define the drop characteristics used by WRED (minimum threshold, maximum threshold, and mark probability denominator). These parameters define the WRED probability slopes.

When a packet arrives at the output queue, the IP precedence or DSCP value is used to select the correct WRED profile for the packet. The packet is then passed to WRED for processing. Based on the selected traffic profile and the average queue length, WRED calculates the probability for dropping the current packet and either drops the packet or passes it to the output queue.

If the queue is already full, the packet is tail-dropped. Otherwise, the packet will eventually be transmitted out onto the interface. If the average queue length is greater than the minimum threshold but less than the maximum threshold, based on the drop probability, WRED will either queue the packet or perform a random drop.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WRED Profiles

This topic describes the different traffic profiles that are used in WRED implementations.

This figure shows two different WRED profiles that are used for traffic of two different QoS classes (“BE” class and “Premium” class).

The BE traffic class has a much lower minimum (10) and maximum threshold (30). As a result, traffic belonging to the BE class will be dropped much earlier and more aggressively than traffic from the Premium class. When heavy congestion occurs, traffic belonging to the BE class will ultimately be tail dropped.

The Premium traffic class has been configured with higher minimum (20) and maximum thresholds (40). Therefore packet drop as a result of congestion will occur later (longer average queue size) and is less likely, compared to the BE class. The differences in these traffic profiles, as defined in the figure, maintain differentiated levels of service in the event of congestion.

To avoid the need for setting all WRED parameters in a router, 8 default values are already defined for precedence-based WRED, and 64 DiffServ aligned values are defined for DSCP-based WRED. Therefore, the default settings should suffice in the vast majority of deployments.

By default, the maximum threshold for all DSCP values is 40. The default mark probability denominator for all DSCP values is 10.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

A PHB is the externally observable forwarding behavior applied at a DiffServ-compliant node to a DiffServ behavior aggregate (BA). With the ability of the system to mark packets according to DSCP setting, collections of packets (each with the same DSCP setting and sent in a particular direction) can be grouped into a DiffServ BA. Packets from multiple sources or applications can belong to the same DiffServ BA.

The class selector BA is used for backward compatibility with non-DiffServ-compliant devices (RFC 1812-compliant devices and, optionally, RFC 791-compliant devices). Therefore, the class selector range of DSCP values is used for backward compatibility with IP precedence. The same WRED profiles are applied to equal IP precedence and class selector values:

IP Precedence and Class Selector Profiles

IP Precedence	DSCP (Class Selector)	Default Minimum Threshold
0 (000)	Default (0)	20
1 (001)	CS1 (8) (001000)	22
2 (010)	CS2 (16) (010000)	24
3 (011)	CS3 (24) (011000)	26
4 (100)	CS4 (32) (100000)	28
5 (101)	CS5 (40) (101000)	30
6 (110)	CS6 (48) (110000)	32
7 (111)	CS7 (56) (111000)	34
RSVP	RSVP	37

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

DSCP-Based WRED (Expedited Forwarding)

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-7

In DSCP, the Expedited Forwarding (EF) PHB is identified based on these parameters:

- Ensures a minimum departure rate to provide the lowest possible delay to delay-sensitive applications.
- Guarantees bandwidth to prevent starvation of the application if there are multiple applications using EF PHB.
- Polices bandwidth to prevent starvation of other applications or classes that are not using this PHB.
- Packets requiring EF should be marked with DSCP binary value “101110” (46 or 0x2E).

For the EF DiffServ traffic class, Cisco IOS DSCP-based WRED configures itself by default so that the minimum threshold is very high, increasing the probability of no drops being applied to that traffic class. It is expected that EF traffic will be dropped very late, compared to other traffic classes, and the EF traffic is therefore prioritized in the event of congestion.

Expedited Forwarding Profile

DSCP (Six Bits)	Default Minimum Threshold
EF (101110)	36

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

In DSCP, the Assured Forwarding (AF) PHB is identified based on these parameters:

- Guarantees a certain amount of bandwidth to an AF class.
- Allows access to extra bandwidth, if available.
- Packets requiring AF PHB should be marked with DSCP value “aaadd0” where “aaa” is the number of the class and “dd” is the drop preference of the traffic class.

There are four defined AF classes. Each class should be treated independently and have bandwidth allocated that is based on the QoS policy. For each AF DiffServ traffic class, Cisco IOS DSCP-based WRED configures itself by default for three different profiles, depending on the drop preference bits. All AF classes are initially marked with Drop Preference 1 (lowest drop preference), but in transit they may be marked down by policers to Drop Preference 2 or Drop Preference 3, depending if they are exceeding or violating administratively defined traffic rates.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Assured Forwarding Profiles

Assured Forwarding Class	Drop Probability	(AF Class) DSCP	Default Minimum Threshold
AF Class 1	Low Drop Prob	(AF11) 001010	32
	Medium Drop Prob	(AF12) 001100	28
	High Drop Prob	(AF13) 001110	24
AF Class 2	Low Drop Prob	(AF21) 010010	32
	Medium Drop Prob	(AF22) 010100	28
	High Drop Prob	(AF23) 010110	24
AF Class 3	Low Drop Prob	(AF31) 011010	32
	Medium Drop Prob	(AF32) 011100	28
	High Drop Prob	(AF33) 011110	24
AF Class 4	Low Drop Prob	(AF41) 100010	32
	Medium Drop Prob	(AF42) 100100	28
	High Drop Prob	(AF43) 100110	24

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring CB-WRED

This topic describes the Cisco IOS commands that are required to configure CB-WRED.

Configuring CB-WRED

```
router(config)#
class-map [match-any | match-all] class-name
```

1. **Create class map: Used for matching packets to a specified class**

```
router(config)#
policy-map policy-name
```

2. **Create policy map (service policy): Specifies a traffic policy that can be attached to one or more interfaces**

```
router(config-if)#
service-policy {input | output} policy-map-name
```

3. **Attach service policy: Associates the policy map to an output interface or VC**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-9

To configure CB-WRED (WRED at the class level with CBWFQ), configure the DSCP-based and precedent-based arguments within MQC. Specific CB-WRED configuration arguments are applied within a policy map. You can then apply the policy map configuration wherever policy maps are attached (for example, at the interface level, the per-virtual circuit [VC] level, or the shaper level).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring CB-WRED (Cont.)

```
router(config-pmap-c)#
random-detect
```

- **Enables IP precedence-based WRED in the selected class within the service policy configuration mode.**
- **Default service profile is used.**
- **Command can be used at the interface, per-VC (with random-detect-group), or the class level (service-policy).**
- **Precedence-based WRED is the default mode.**
- **WRED treats non-IP traffic as precedence 0.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-10

Use the **random-detect** command to enable WRED on an interface. By default, WRED is precedence-based and uses eight default WRED profiles, one for each value of IP precedence.

Within the CBWFQ system, WRED is used to perform per-queue dropping within the class queues. Therefore, each class queue has its own WRED method, which can be further weighed based on the IP precedence or DSCP value. Each queue can therefore be configured with a separate dropping policy to implement different drop policies for every class of traffic.

WRED will treat all non-IP traffic as precedence 0. As a result, non-IP traffic is more likely to be dropped than IP traffic.

If the **random-detect** command is used on versatile interface processor (VIP)-based interfaces, distributed WRED (DWRED) is enabled and the VIP CPU is responsible for WRED dropping. This can significantly increase router performance when used in the context of distributed Cisco Express Forwarding (CEF) switching, which is a prerequisite for DWRED functionality. Also, DWRED can be combined with distributed weighted fair queuing (DWFQ), enabling truly distributed queuing and congestion avoidance techniques, running independently from the central CPU.

WRED cannot be configured on the same interface as custom queuing (CQ), priority queuing (PQ), or WFQ. However, both DWRED and DWFQ can be configured on the same interface. In addition, CB-WRED can be configured in conjunction with CBWFQ. Restricting nondistributed, non-class-based WRED to only FIFO queuing on an interface is typically not a major issue because WRED is usually applied in the network core, where advanced queuing mechanisms are not typically used. WRED is suited for the network core because WRED has a relatively low performance impact on routers. Furthermore, DWRED or CB-WRED can be used to overcome this limitation by combining WRED with WFQ.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Changing the WRED Traffic Profile

```
router(config-pmap-c)#
random-detect precedence precedence min-threshold
max-threshold mark-prob-denominator
```

- **Changes WRED profile for specified IP precedence value.**
- **Packet drop probability at maximum threshold is:**
  1 / mark-prob-denominator
- **Non-weighted RED is achieved by using the same WRED profile for all precedence values.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-11

When WRED is enabled, default values are selected for each traffic profile based on the weight used (IP precedence or DSCP). You can then modify these default values to match their specific administrative QoS policy goals. When modifying the default WRED profile for IP precedence, these values are configurable:

- **Minimum threshold:** When the average queue depth is above the minimum threshold, WRED starts dropping packets. The rate of packet drop increases linearly as the average queue size increases, until the average queue size reaches the maximum threshold. The default maximum threshold is equal to the default hold queue size (40) on an interface. The size of the hold queue is equivalent to the number of packets that can be held within a queue. The hold queue length ranges from 0 to 4096, and, therefore, the minimum/maximum threshold range is 1 to 4096. The default maximum threshold will reflect the defined hold queue size. Thus, if the hold queue is changed, the maximum threshold will change.
- **Maximum threshold:** When the average queue size is above the maximum threshold, all packets are dropped. If the difference between the maximum threshold and the minimum threshold is too small, many packets might be dropped at once, resulting in global synchronization.
- **Mark probability denominator:** This is the fraction of packets dropped when the average queue depth is at the maximum threshold. For example, if the denominator is 10, one out of every 10 packets is dropped when the average queue is at the maximum threshold. The maximum probability of drop at the maximum threshold can be expressed as 1 / mark-prob-denominator. The maximum drop probability is 10 percent if default settings are used that have a mark probability denominator value of 10. The value of the mark probability can range from 1 to 65536.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

If required, you can configure RED as a special case of WRED, by assigning the same profile to all eight IP precedence values. The default WRED parameter is based on the best available data. Cisco recommends that these parameters not be changed from their default values unless you have determined that your applications will benefit from the changed values.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Changing WRED Sensitivity to Bursts

```
router(config-pmap-c)#
random-detect exponential-weighting-constant n
```

- **WRED takes the average queue size to determine the current WRED mode (no drop, random drop, full drop).**

$$Q_{avg}(t+1) = Q_{avg}(t) \cdot (1 - 2^{-n}) + Q_t \cdot 2^{-n}$$

New average queue size — Previous average queue size — Current queue size

017G_355

- **High values of N allow short bursts.**
- **Low values of N make WRED more burst-sensitive.**
- **Default value (9) should be used in most scenarios.**
- **Average output queue size with N=9 is**

$$Q_{ave}(t+1) = Q_{ave}(t) * 0.998 + Q_t * 0.002$$

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-12

WRED does not calculate the drop probability using the current queue length, but rather uses the average queue length. The average queue length is constantly recalculated using two terms: the previously calculated average queue size and the current queue size. An exponential weighting constant N influences the calculation by weighing the two terms, therefore influencing how the average queue size follows the current queue size, in this way:

- For high values of N, the previous average becomes more important. A large factor will smooth out the peaks and lows in queue length. The average queue size is unlikely to change very quickly, avoiding drastic swings in size. The WRED process will be slow to start dropping packets, but it may continue dropping packets for a time after the actual queue size has fallen below the minimum threshold. The slow-moving average will accommodate temporary bursts in traffic.

Note If the value of N gets too high, WRED will not react to congestion. Packets will be transmitted or dropped as if WRED were not in effect.

- For low values of N, the average queue size closely tracks the current queue size. The resulting average may fluctuate with changes in the traffic levels. In this case, the WRED process responds quickly to long queues. When the queue falls below the minimum threshold, the process will stop dropping packets.

Note If the value of N gets too low, WRED will overreact to temporary traffic bursts and drop traffic unnecessarily.

The default value of N is 9. This value should suffice for most scenarios, except perhaps those involving extremely high-speed interfaces (such as OC12), where the value can be increased slightly (to about 12) to allow more bursts.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: CBWFQ Using IP Precedence with CB-WRED

Example: CBWFQ Using IP Precedence with CB-WRED

- **Enable CBWFQ to prioritize traffic according to these requirements:**
  - **Class mission-critical is marked with IP precedence values 3 and 4 (3 is high-drop, 4 is low-drop) and should get 30% of interface bandwidth.**
  - **Class bulk is marked with IP precedence values 1 and 2 (1 is high-drop, 2 is low-drop) and should get 20% of interface bandwidth.**
  - **All other traffic should be per-flow fair-queued.**
- **Use differentiated WRED to prevent congestion in all three classes.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-13

This example of CBWFQ with WRED focuses on a network that provides these three different service levels for three traffic classes:

- **Mission-critical class:** Marked with IP Precedence values 3 and 4 (3 is used for high-drop service, and 4 is used for low-drop service within the service class), and should get 30 percent of an interface bandwidth
- **Bulk class:** Marked with IP Precedence values 1 and 2 (1 is used for high-drop service, and 2 is used for low-drop service), and should get 20 percent of the interface bandwidth
- **Best-effort class:** Should get the remaining bandwidth share, and should be fair-queued

To enforce this service policy, a router will use CBWFQ to perform bandwidth sharing and WRED within service classes to perform differentiated drop.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: CBWFQ Using IP Precedence with CB-WRED

```
class-map Mission-critical
 match ip precedence 3 4
!
class-map Bulk
 match ip precedence 1 2
!
policy-map Policy1
  class Mission-critical
   bandwidth percent 30
   random-detect
   random-detect precedence 3 26 40 10
   random-detect precedence 4 28 40 10
```

```
class Bulk
   bandwidth percent 20
   random-detect
   random-detect precedence 1 22 36 10
   random-detect precedence 2 24 36 10
  class class-default
   fair-queue
   random-detect
!
```

017G_353

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-14

The figure shows the WRED traffic profile representing the QoS service policy, and the configuration that is used to implement the example service policy. The traffic is classified based on the precedence bits, and all noncontract traffic is classified into the default class, as follows:

- The mission-critical class is guaranteed at least 30 percent of bandwidth with a custom WRED profile that establishes a low-drop and a high-drop per-hop behavior.
- The bulk class is guaranteed at least 20 percent of bandwidth, is configured with somewhat lower WRED drop thresholds, and is therefore more likely to be dropped than the mission-critical class in the event of interface congestion.
- All other traffic is part of the default class and is fair-queued with default WRED parameters.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring DSCP-Based CB-WRED

This topic describes the Cisco IOS commands that are required to configure DSCP-based CB-WRED.

Configuring DSCP-Based CB-WRED

```
router(config-pmap-c)#
random-detect dscp-based
```

- **Enables DSCP-based WRED.**
- **Command can be used at the interface, per-VC (with random-detect-group), or the class level (service-policy).**
- **Default service profile is used.**
- **The WRED** random-detect **command and the WFQ** queue-limit **command are mutually exclusive for class policy.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-15

Use the **random-detect dscp-based** command to enable DSCP-based WRED on an interface. By default, WRED is precedence-based and uses eight default WRED profiles, one for each value of IP precedence. Changing WRED weighting to values based on DSCP increases the number of WRED traffic profiles to 64.

You can configure WRED as part of the policy for a standard class or the default class. The WRED **random-detect** command and the WFQ **queue-limit** command are mutually exclusive for class policy. If you configure WRED, its packet-drop capability is used to manage the queue when packets exceeding the configured maximum count are enqueued. If you configure the WFQ **queue-limit** command for class policy, tail drop is used.

WRED cannot be configured on the same interface as CQ, PQ, or WFQ. However, both DWRED and DWFQ can be configured on the same interface. In addition, CB-WRED can be configured in conjunction with CBWFQ. Restricting nondistributed, non-class-based WRED only to FIFO queuing on an interface is not a major issue because WRED is usually applied in the network core, where advanced queuing mechanisms are not typically deployed. WRED is suited for the network core as it has a relatively low performance impact on routers. Further, DWRED or CB-WRED can be used to overcome this limitation by combining WRED with WFQ.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Changing the WRED Traffic Profile

```
router(config-pmap-c)#
random-detect dscp dscpvalue min-threshold max-
threshold mark-prob-denominator
```

- **Changes WRED profile for specified DSCP value**
- **Packet drop probability at maximum threshold is:**
  1 / mark-prob-denominator

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-16

When DSCP-based WRED is enabled, default values are selected for each traffic profile based on DSCP. You can modify these default values to match their specific administrative QoS policy goals. When modifying the default WRED profile for DSCP, these values are configurable:

- **Minimum threshold:** When the average queue depth is above the minimum threshold, WRED starts dropping packets. The rate of packet drop increases linearly as the average queue size increases, until the average queue size reaches the maximum threshold. The default maximum threshold is equal to the default hold queue size (40) on an interface. The size of the hold queue is equivalent to the number of packets that can be held within a queue. The hold queue length ranges from 0 to 4096, and, therefore, the minimum/maximum threshold range is 1 to 4096. The default maximum threshold will reflect the defined hold queue size. Thus, if the hold queue is changed, the maximum threshold will change.
- **Maximum threshold:** When the average queue size is above the maximum threshold, all packets are dropped. If the difference between the maximum threshold and the minimum threshold is too small, many packets might be dropped at once, resulting in global synchronization.
- **Mark probability denominator:** This is the fraction of packets dropped when the average queue depth is at the maximum threshold. For example, if the denominator is 10, one out of every 10 packets is dropped when the average queue is at the maximum threshold. The maximum probability of drop at the maximum threshold can be expressed as 1 / mark-prob-denominator. The maximum drop probability is 10 percent if default settings are used that have a mark probability denominator value of 10. The value of the mark probability can range from 1 to 65536.

Note The default WRED parameters are based on the best available data. Cisco recommends that these parameters not be changed from their default values.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: CB-WRED Using DSCP with CBWFQ

In this example of CBWFQ with WRED, the service policy was implemented using classes of service based on IP precedence.

Example: CB-WRED Using DSCP with CBWFQ

- **Enable CBWFQ to prioritize traffic according to these requirements:**
  - **Class mission-critical is marked using DSCP AF2 and should get 30% of interface bandwidth.**
  - **Class bulk is marked using DSCP AF1 and should get 20% of interface bandwidth.**
  - **All other traffic should be per-flow fair-queued.**
- **Use differentiated WRED to prevent congestion in all three classes.**
- **Make sure the new configurations still conform to the design and implementation from the previous example.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-17

The same service policy will be configured in this example. In this case, however, DSCP-based classes of service are used. Remember that the DiffServ model itself provides defined traffic classes and their associated PHB. DiffServ-based classification is used in this example as follows:

- **Mission-critical class:** Marked using DSCP AF class 2, and should get 30 percent of an interface bandwidth
- **Bulk class:** Marked using DSCP AF class 1, and should get 20 percent of the interface bandwidth
- **Best-effort class:** Traffic should get the remaining bandwidth share, and should be fair-queued

To enforce this service policy, a router will use CBWFQ to perform bandwidth sharing, and will use WRED within service classes to perform differentiated drop.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: CB-WRED Using DSCP with CBWFQ (Cont.)

```
class-map Mission-critical
 match ip dscp af21 af22 af23 cs2
!
class-map Bulk
 match ip dscp af11 af12 af13 cs1
!
policy-map Policy1
  class Mission-critical
   bandwidth percent 30
   random-detect dscp-based
   random-detect dscp af21 32 40 10
   random-detect dscp af22 28 40 10
   random-detect dscp af23 24 40 10
   random-detect dscp cs2 24 40 10
```

```
 class Bulk
   bandwidth percent 20
   random-detect dscp-based
   random-detect dscp af11 32 36 10
   random-detect dscp af12 28 36 10
   random-detect dscp af13 24 36 10
   random-detect dscp cs1 22 36 10
 class class-default
   fair-queue
   random-detect dscp-based
!
```

017G_354

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-18

The configuration example shows how traffic classification is performed using DSCP-based classes, representing the mission-critical class as the AF1 class, and the AF2 class as the bulk class. WRED DSCP-based parameters are set reflecting the class-dependent drop strategy, as follows:

- The mission-critical class is guaranteed at least 30 percent of bandwidth, with a custom WRED profile which establishes three different drop probabilities for AF class 2.
- The bulk class is guaranteed at least 20 percent of bandwidth, is configured with three different drop probabilities for AF class 1, and has a somewhat lower WRED maximum threshold. As a result, bulk-class traffic is more likely to be dropped than the mission-critical class in the event of interface congestion.

All other traffic is part of the default class, and is fair-queued, with default WRED parameters.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring CB-WRED

This topic describes the Cisco IOS commands that are required to monitor CB-WRED.

Monitoring CB-WRED

```
router#
show policy-map interface interface-name
```

- **Displays the configuration of all classes configured for all service policies on the specified interface**

```
router#show policy-map interface Ethernet 0/0
 Ethernet0/0
  Service-policy output: Policy1
    Class-map: Mission-critical (match-all)
      0 packets, 0 bytes  5 minute offered rate 0 bps, drop rate 0 bps
      Match: ip precedence 2  Match: ip dscp 18  20  22
      Weighted Fair Queueing
        Output Queue: Conversation 265
        Bandwidth 30 (%)  Bandwidth 3000 (kbps)
        (pkts matched/bytes matched) 0/0
        (depth/total drops/no-buffer drops) 0/0/0
        exponential weight: 9
        mean queue depth: 0
    Dscp     Transmitted Random drop Tail drop   Minimum   Maximum     Mark
    (Prec)   pkts/bytes  pkts/bytes  pkts/bytes  threshold threshold probability
      0(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
...
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-19

Use the **show policy-map interface** command to display the configuration of all classes configured for all service policies on the specified interface. This includes all WRED parameters implementing the drop policy on the specified interface.

The table explains some of the key fields of the output of the **show policy-map interface** command.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

show policy-map interface Output Fields

Field	Description
`service-policy output`	Name of the output service policy applied to the specified interface or VC.
`class-map`	Class of traffic being displayed. Output is displayed for each configured class in the policy. The choice for implementing class matches (for example, **match-all** or **match-any**) can also appear next to the traffic class.
`match`	Match criteria specified for the class of traffic. Choices include criteria such as IP precedence, IP DSCP value, Multiprotocol Label Switching (MPLS) experimental value, access groups, and QoS groups.
`exponential weight`	Exponent used in the average queue size calculation for a WRED parameter group.
`mean queue depth`	Average queue depth based on the actual queue depth on the interface and the exponential weighting constant. It is a fluctuating average. The minimum and maximum thresholds are compared against this value to determine drop decisions.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **WRED combines RED with IP precedence or DSCP and performs packet dropping based on IP precedence or DSCP markings.**
- **Each WRED profile defines the minimum and maximum threshold and the maximum drop probability. Profiles are already defined by default for IP precedence and DSCP.**
- **The** random-detect **command is used to enable WRED on an interface.**
- **The** random-detect dscp-based **command is used to enable DSCP-based WRED on an interface.**
- **The** show policy-map interface **command displays the configuration of all classes configured for all service policies on the specified interface. This includes all WRED parameters implementing the drop policy on the specified interface.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-20

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 4

Configuring Explicit Congestion Notification

Overview

Congestion avoidance techniques monitor network traffic loads in an effort to anticipate and avoid congestion at common network bottleneck points. You achieve congestion avoidance through packet dropping, by using more complex techniques than simple tail drop. With the addition of explicit congestion notification (ECN) extensions to IP, routers have an alternative method of indicating congestion to peers.

This lesson describes the concept of ECN and the Cisco IOS commands that are required to configure and monitor ECN.

Objectives

Upon completing this lesson, you will be able to configure ECN to enhance the congestion avoidance features of WRED. This ability includes being able to meet these objectives:

- Describe the ECN extensions to IP
- Identify characteristics of the ECN field in IP
- Explain how ECN interacts with WRED
- Identify the Cisco IOS commands that are required to configure ECN extensions to WRED
- Identify the Cisco IOS commands that are required to monitor ECN-enabled WRED

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Explicit Congestion Notification

This topic describes the ECN extensions to IP.

Explicit Congestion Notification

- **TCP congestion controls are not suited to applications that are sensitive to delay or packet loss.**
- **ECN (RFC 3168) removes need to rely on packet loss as a congestion indicator.**
- **ECN marks packets instead of dropping them when the average queue length exceeds a specific threshold value.**
- **Routers and end hosts can use ECN marking as a signal that the network is congested and send packets at a slower rate.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-2

TCP determines how many unacknowledged packets it can send by gradually increasing the number of packets the session sends until it experiences a dropped packet; this is known as the TCP window size. As a result, TCP tends to cause router queues to build up at network bottleneck points. When queues become full, tail drop begins dropping all incoming packets until there is room in the queue. Tail drop does not provide differential treatment and therefore some of the fragile flow packets, sensitive to latency, may be dropped. In addition, tail drop can lead to global synchronization of packet loss across multiple flows.

Active queue management mechanisms such as random early detection (RED) or weighted random early detection (WRED) detect congestion before queues fill and overflow. Through the use of selective packet discard, these mechanisms provide congestion indication to end nodes. Therefore, active queue management (congestion avoidance) mechanisms can reduce queuing delays for all traffic sharing a specific queue. In addition, active queue management means that it is no longer necessary to rely on buffer overflow as the only means of indicating congestion.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Traditional active queue management mechanisms, such as RED, rely on the use of packet drops to indicate congestion. Packet dropping in these mechanisms is based on the average queue length exceeding a predefined threshold, rather than only when queues overflow. However, because packets are dropped prior to queues actually overflowing, the router dropping the packet is not always constrained by memory limitations and needs to actually drop the packet. With the *Addition of Explicit Congestion Notification to IP* (RFC 3168), active queue management allows routers to signal that congestion has been experienced by the router, instead of relying on the use of packet drops. Through the use of signaling congestion, aggressive flows can be slowed, thus reducing the impact of congestion and packet loss on latency-sensitive flows.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

ECN Field Defined

This topic describes the characteristics of the ECN field in IP.

The Addition of Explicit Congestion Notification to IP (RFC 3168) redefines the DiffServ field (formerly the type of service [ToS] byte) to contain an ECN-specific field. The ECN field consists of the last two low-order bits of the DiffServ field, and consists of the ECN-capable transport (ECT) bit and the congestion experienced (CE) bit.

The ECT bit and the CE bit can be used to make four ECN field combinations: 00, 01, 10, and 11. The different ECT and CE bit combinations in the ECN field have these meanings:

- **00:** The ECN field combination indicates that a packet is not using ECN.
- **01 and 10:** The ECN field combinations, called ECT(1) and ECT(0), respectively, are set by the data sender to indicate that the endpoints of the transport protocol are ECN-capable. Routers will treat these two field combinations identically. Data senders can use either one or both of these two combinations.
- **11:** The ECN field combination indicates to the endpoints that congestion has been experienced. Packets arriving at a full queue of a router will be dropped.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

ECN and WRED

This topic describes how ECN interacts with WRED.

ECN and WRED

- **ECN is an extension to WRED.**
- **Congestion in WRED is indicated based on the average queue length exceeding a specific threshold value.**
- **If the number of packets in the queue is below the minimum threshold, packets are transmitted.**
  - **Treatment is identical to a network using only WRED.**
- **If the number of packets in the queue is above the maximum threshold, packets are tail-dropped.**
  - **Treatment is identical to a network using only WRED.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-4

On Cisco IOS routers, ECN is an extension to WRED functionality. WRED is an active queue management mechanism that uses packet drops as a congestion indicator to endpoints. Packets are dropped by WRED based on the average queue length exceeding a specific set of predefined threshold values (minimum and maximum threshold). ECN is an extension to WRED, in that ECN marks packets instead of dropping them when the average queue length exceeds a specific threshold value. When ECN is configured with WRED, routers and end hosts use this marking as a signal that the network is congested and slow down the rate at which packets are sent.

ECN must be interoperable with non-ECN-compliant devices. Because ECN is configured as an extension to WRED, packets are treated differently by WRED when ECN has been enabled.

If the average queue length is below the defined WRED minimum threshold, all packets are queued and transmitted normally. This behavior is identical to devices that are configured to use non-ECN-enabled WRED.

If the average queue length is greater than the maximum threshold, packets are tail-dropped. This behavior is identical to devices configured to use non-ECN-enabled WRED.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

ECN and WRED (Cont.)

If the number of packets in the queue is between the minimum and maximum threshold, one of three scenarios can occur:

- **ECN-capable endpoints and WRED determine that the packet should be dropped based on the drop probability:**
  - **ECT and CE bits for the packet are changed to 1 and the packet is transmitted.**
- **Non ECN-capable endpoints:**
  - **The packet may be dropped based on the WRED drop probability.**
- **The network is experiencing congestion:**
  - **Both the ECT and CE are already set. The packet is transmitted and no further marking is required.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-5

If the number of packets in the queue is below the minimum threshold, packets are transmitted. This happens whether or not ECN is enabled, and this treatment is identical to the treatment a packet receives when WRED only is being used on the network.

If the number of packets in the queue is between the minimum threshold and the maximum threshold, one of these three scenarios can occur:

- If the ECN field on the packet indicates that the endpoints are ECN-capable (that is, the ECT bit is set to 1 and the CE bit is set to 0, or the ECT bit is set to 0 and the CE bit is set to 1) and the WRED algorithm determines that the packet should have been dropped based on the drop probability, the ECT and CE bits for the packet are changed to 1 and the packet is transmitted. This happens because ECN is enabled and the packet gets marked instead of dropped.
- If the ECN field on the packet indicates that neither endpoint is ECN-capable (that is, the ECT bit is set to 0 and the CE bit is set to 0), the packet may be dropped based on the WRED drop probability. This is the identical treatment that a packet receives when WRED is enabled without ECN configured on the router.
- If the ECN field on the packet indicates that the network is experiencing congestion (that is, both the ECT bit and the CE bit are set to 1), the packet is transmitted. No further marking is required.

If the number of packets in the queue is above the maximum threshold, packets are dropped based on the drop probability. This is the identical treatment a packet receives when WRED is enabled without ECN configured on the router.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring ECN-Enabled WRED

This topic describes the Cisco IOS commands that are required to configure ECN extensions to WRED.

Configuring ECN-Enabled WRED

```
router(config-pmap-c)#
random-detect ecn
```

- **Enables ECN.**
- **ECN can be used whether WRED is based on the IP precedence or DSCP value.**
- **ECN must be configured through MQC.**

```
router(config)# policy-map MyPolicy
router(config-pmap)# class class-default
router(config-pmap-c)# bandwidth percent 70
router(config-pmap-c)# random-detect
router(config-pmap-c)# random-detect ecn
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-6

The ECN feature must be configured through the Modular QoS command-line interface (CLI), or MQC. ECN is configured as part of a policy map after CB-WRED has been enabled. ECN can be used whether the CB-WRED configuration is based on IP precedence or differentiated services code point (DSCP).

Note The WRED ECN feature was introduced in Cisco IOS Release 12.2(8)T.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring ECN-Enabled WRED

This topic describes the Cisco IOS commands that are required to monitor ECN-enabled WRED.

Monitoring ECN-Enabled WRED

```
router#
show policy-map [policy-map]
```

- **Displays the configuration of all classes for a specified service policy map or all classes for all existing policy maps**

```
router#show policy-map
Policy Map MyPolicy
Class class-default
Weighted Fair Queueing
Bandwidth 70 (%)
exponential weight 9
explicit congestion notification
class min-threshold max-threshold mark-probability
------------------------------------------------------------
------------------------------------------------------------
0         -         -         1/10
1         -         -         1/10
2         -         -         1/10
3         -         -         1/10
. . .
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-7

The **show policy-map** command displays the configuration of a service policy map created using the **show policy-map** command. The **show policy-map** command will display ECN marking information only if ECN is enabled on the interface. The table explains some of the key fields of the output of the **show policy-map** command.

show policy-map Output Fields

Field	Description
`explicit congestion notification`	Indication that ECN is enabled
`class`	IP precedence value
`min-threshold`	Minimum WRED threshold in number of packets
`max-threshold`	Maximum WRED threshold in number of packets
`mark-probability`	Fraction of packets dropped when the average queue depth is at the maximum threshold

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring ECN-Enabled WRED (Cont.)

```
router#
show policy-map interface interface-name
```

- **Displays the configuration of all classes configured for all service policies on the specified interface**

```
router#show policy-map interface Serial4/1
Serial4/1
Service-policy output:policy_ecn
Class-map:prec1 (match-all)
1000 packets, 125000 bytes
30 second offered rate 14000 bps, drop rate 5000 bps
Match:ip precedence 1
Weighted Fair Queueing
Output Queue:Conversation 42
Bandwidth 20 (%)
Bandwidth 100 (kbps)
(pkts matched/bytes matched) 989/123625
(depth/total drops/no-buffer drops) 0/455/0
exponential weight:9
explicit congestion notification
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-8

The **show policy-map interface** command displays the configuration of all classes configured for all service policies on the specified interface. The counters displayed after you enter the **show policy-map interface** command are updated only if congestion is present on the interface. The **show policy-map interface** command also displays ECN marking information only if ECN is enabled on the interface. The table explains some of the key fields of the output of the **show policy-map interface** command.

show policy-map interface Output Fields

Field	Description
`explicit congestion notification`	Indication that ECN is enabled

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring ECN-Enabled WRED (Cont.)

```
mean queue depth:0
class Transmitted Random drop Tail drop  Minimum   Maximum   Mark
      pkts/bytes  pkts/bytes  pkts/bytes threshold threshold probability
0         0/0       0/0       0/0          20        40        1/10
1         545/68125 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

     class        ECN Mark
                 pkts/bytes
       0            0/0
       1            43/5375
       2            0/0
       3            0/0
       4            0/0
       5            0/0
       6            0/0
       7            0/0
     rsvp           0/0
```

© 2006 Cisco Systems, Inc. All rights reserved.

show policy-map interface Output Fields (Cont.)

Field	Description
`mean queue depth`	Average queue depth based on the actual queue depth on the interface and the exponential weighting constant. It is a moving average. The minimum and maximum thresholds are compared against this value to determine drop decisions.
`class`	IP precedence value.
`transmitted pkts/bytes`	Number of packets (also shown in bytes) transmitted.
`random drop pkts/bytes`	Number of packets (also shown in bytes) randomly dropped when the mean queue depth is between the minimum threshold value and the maximum threshold value for the specified IP precedence value.
`tail drop pkts/bytes`	Number of packets dropped when the mean queue depth is greater than the maximum threshold value for the specified IP precedence value.
`minimum threshold`	Minimum WRED threshold in number of packets.
`maximum threshold`	Maximum WRED threshold in number of packets.
`mark probability`	Fraction of packets dropped when the average queue depth is at the maximum threshold.
`ECN mark pkts/bytes`	Number of packets (also shown in bytes) marked by ECN.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **ECN marks packets instead of dropping them when the average queue length exceeds a specific threshold value.**
- **ECN defines two flow control bits as extensions to the DiffServ field: The ECT bit and the CE bit.**
- **ECN is an extension to WRED that removes the need to rely on packet loss as a congestion indicator.**
- **On Cisco IOS routers, the ECN feature must be configured through MQC. Use the** random-detect ecn **command.**
- **The** show policy-map interface **command will display ECN marking information only if ECN is enabled on the interface.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-10

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Summary

This topic summarizes the key points discussed in this module.

Module Summary

- **By default, routers use tail drop as a means of congestion control when an output queue is full. Tail drop treats all traffic equally and does not differentiate between classes of service. When tail drop is in effect, packets are dropped until the congestion is eliminated and the queue is no longer full.**
- **Congestion avoidance techniques, such as RED, monitor network traffic loads in an effort to anticipate and avoid congestion at common network bottlenecks. Congestion avoidance is achieved through packet dropping.**
- **WRED, the Cisco implementation of RED, combines the capabilities of the RED algorithm with IP precedence or DSCP.**
- **ECN is an extension to WRED that enables flow control and congestion signaling without requiring packet drops.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—6-1

Congestion management is an area of concern for all networks that require a differentiated treatment of packet flows. Active queue management mechanisms address the limitations of relying solely on TCP congestion management techniques, which simply wait for queues to overflow and then drop packets to signal that congestion has occurred.

Congestion avoidance mechanisms such as RED and WRED allow for specific packet flows to be selectively penalized and slowed by applying a traffic profile. Traffic flows are matched against this profile and transmitted or dropped, depending upon the average length of the interface output queue. In addition, RED and WRED are extremely effective tools at preventing global synchronization of many TCP traffic flows.

Another active queue management technique is ECN. ECN is an extension to WRED that allows for signaling to be sent to ECN-enabled endpoints, instructing them to reduce their packet transmission rates. ECN also provides the benefit of not requiring packet drops when the WRED drop probability indicates otherwise.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

References

For additional information, refer to these resources:

- Further details on TCP slow start can be found in RFC 2001 at this URL: http://www.faqs.org/rfcs/rfc2001.html
- For more information, and a list of operating systems supporting RFC 1323, refer to: http://www.psc.edu/networking/perf_tune.html#table.
- For a detailed discussion of TCP protocol behavior see Geoff Huston, Telstra, "TCP Performance," *Internet Protocol Journal*, Vol. 3, No. 2, June 2000, at this URL: http://www.cisco.com/warp/public/759/ipj_3-2/ipj_3-2_tcp.html
- For a detailed discussion of TCP congestion behavior see Geoff Huston, Telstra, "The Future for TCP," *Internet Protocol Journal*, Vol. 3, No. 3, September 2000, at this URL: http://www.cisco.com/warp/public/759/ipj_3-3/ipj_3-3_futureTCP.html
- For more information on Random Early Detection, refer to, *Congestion Avoidance Overview* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121cgcr/qos_c/qcprt3/qcdconav.htm
- For more information on WRED and configuring WRED, refer to *Configuring Weighted Random Early Detection* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/qos_c/qcpart3/qcwred.htm
- For more information on DSCP-based WRED, refer to *DiffServ Compliant Weighted Random Early Detection* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121t/121t5/dtdswred.htm
- For information regarding WRED on Cisco GSR 12000 routers, refer to *Weighted Random Early Detection on the Cisco 12000 Series Router* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios112/ios112p/gsr/wred_gs.htm
- For more information on ECN, refer to *RFC 3168: The Addition of Explicit Congestion Notification to IP* at this URL: http://www.ietf.org/rfc/rfc3168.txt
- For more information on using WRED with ECN, refer to *WRED—Explicit Congestion Notification* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios124/124cg/hqos_c/part15/ch05/hwrdecn.htm

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check

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

Q1) What are two ways in which TCP manages congestion? (Choose two.) (Source: Introducing Congestion Avoidance)

A) TCP uses tail drop on queues that have reached their queue limit.
B) TCP uses dropped packets as an indication that congestion has occurred.
C) TCP uses variable window sizes to reduce and increase the rates at which packets are sent.
D) TCP measures the average size of device queues and drops packets, linearly increasing the amount of dropped packets with the size of the queue.

Q2) Two stations (A and B) are communicating using TCP. Station A has negotiated a TCP window size of 5 and as a result sends 5 packets to station B.

Station A receives 3 ACK messages from station B indicating ACK 3.

Which two of the following descriptions best describe the status of the communication between A and B? (Choose two.) (Source: Introducing Congestion Avoidance)

A) Station B is acknowledging receipt of packets 1, 2, and 3, but has lost packets 4 and 5.
B) Station A initiates a fast-retransmit and immediately sends packet 3 to B.
C) Station B has not received packet 3.
D) Station B has received packets 1, 2, and 3, but not packet 4. It cannot be determined where packet 5 was received at B until packet 4 has been sent.
E) Station A will send packets 4 and 5 to station B upon receipt of the station B ACK.

Q3) What are three important limitations of using a tail-drop mechanism to manage queue congestion? (Choose three.) (Source: Introducing Congestion Avoidance)

A) Tail drop can cause many flows to synchronize, lowering overall link utilization.
B) Tail drop can cause starvation of fragile flows.
C) Tail drop increases the amount of packet buffer memory required, because queues must be full before congestion management becomes active.
D) Tail drop results in variable delays, which can interfere with delay-sensitive traffic flows.

Q4) What are three advantages of active congestion management using RED? (Choose three.) (Source: Introducing RED)

A) RED uses selective packet discard to eliminate global synchronization of TCP flows.
B) RED avoids congestion by ensuring that interface queues never become full.
C) RED increases the overall utilization of links.
D) RED uses selective packet discard to penalize aggressive flows.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q5) A specific RED profile has been configured with a mark probability denominator of 1. What is the effect of this configuration on packet loss as the average queue length reaches the maximum threshold? (Source: Introducing RED)

A) Given this configuration, no packets will be dropped until the average queue length is greater than the maximum threshold.
B) For every active traffic flow, one packet will be discarded.
C) When the average queue length is at the maximum threshold, all packets are dropped.
D) This is an invalid configuration.

Q6) Refer to the following RED traffic profile. In which two ways will the RED traffic profile in this figure affect the traffic flows to which it is applied? (Choose two.) (Source: Introducing RED)

Figure: RED Traffic Profile

A) Packets may be dropped unnecessarily because the minimum threshold is too low.
B) This profile can result in global synchronization because the difference between the minimum and maximum thresholds is too small.
C) RED will not be effective because the mark probability denominator is 50 when the average queue length reaches the maximum threshold.
D) The reduced size of the maximum threshold will prevent tail drop and maximize link utilization.

Q7) What are the three traffic drop modes in random early detection? (Choose three.) (Source: Introducing RED)

A) no drop
B) full drop
C) random drop
D) deferred drop

Q8) Which two QoS markers can you base the “weight” in WRED on when configuring CB-WRED? (Choose two.) (Source: Configuring Class-Based Weighted RED)

A) CoS
B) DSCP
C) QoS group
D) IP precedence

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q9) What are two requirements for configuring CB-WRED? (Choose two.) (Source: Configuring Class-Based Weighted RED)

A) An MQC configuration that includes a policy map must be configured.
B) CEF must be enabled for IP.
C) Random detect must be enabled for DSCP-based CB-WRED.
D) A previous configuration of CBWFQ must be present.

Q10) Given the following CB-WRED configuration, which command should be entered in the Bronze traffic class to properly enable CB-WRED using a minimum threshold of 22, a maximum threshold of 36, and a drop probability of 10 percent? (Source: Configuring Class-Based Weighted RED)

```
class-map Bronze
 match ip dscp cs1
!
policy-map Policy1
  class Bronze
    bandwidth percent 15
    random-detect dscp-based
[ <- _____ -> ]
!
  class class-default
    fair-queue
    random-detect dscp-based
```

A) **random-detect dscp-based 22 36 10**
B) **random-detect dscp-based cs1 22 36 10**
C) **random-detect dscp cs1 22 36**
D) **random-detect dscp cs1 10 22 36**

Q11) What are two active congestion management mechanisms available on Cisco IOS routers? (Choose two.) (Source: Configuring Class-Based Weighted RED, Configuring Explicit Congestion Notification)

A) tail drop
B) weighted round robin
C) explicit congestion notification
D) weighted random early detection

Q12) What will a router do with a newly arriving packet if its output queue is full and ECN fields are both set to a 1? (Source: Configuring Explicit Congestion Notification)

A) drop the last packet on the queue and enqueue the newly arriving packet
B) perform a tail drop and drop the new packet
C) move the packet to the head of the queue to ensure that the receiver is signaled about the network congestion condition
D) allocate additional interface buffers to store the packet because it contains congestion notification information

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q13) Given the following configuration, which Cisco IOS configuration command must be added to the default class to enable ECN? (Source: Configuring Explicit Congestion Notification)

```
policy-map MyPolicy
  class class-default
  bandwidth percent 70
  random-detect
[ <- _____ -> ]
```

A) **wred ecn**
B) **ecn enable**
C) **random-detect ecn**
D) **random-detect ecn enable**

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check Answer Key

Q1)	B, C
Q2)	B, C
Q3)	A, B, D
Q4)	A, C, D
Q5)	C
Q6)	A, B
Q7)	A, B, C
Q8)	B, D
Q9)	A, D
Q10)	C
Q11)	C, D
Q12)	B
Q13)	C

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module 7

Traffic Policing and Shaping

Overview

Within a network, different forms of connectivity can have significantly different costs to an organization. Because WAN bandwidth is relatively expensive, many organizations would like to limit the amount of traffic that specific applications can send. This is especially true when enterprise networks use Internet connections for remote site and extranet connectivity. Downloading non-business-critical images, music, and movie files can greatly reduce the amount of bandwidth available to other mission-critical applications. Traffic policing and traffic shaping are two quality of service (QoS) techniques that can be used to limit the amount of bandwidth that a specific application can use on a link.

This module describes the operation of traffic policing and traffic shaping and how these techniques can be used to rate-limit traffic. Frame Relay WANs have specific requirements; therefore, class-based traffic shaping on Frame Relay networks is also covered in this module.

Module Objectives

Upon completing this module, you will be able to use Cisco QoS traffic policing and traffic-shaping mechanisms to effectively limit the rate of network traffic. This ability includes being able to meet these objectives:

- Explain how traffic policing and traffic shaping can be used to condition traffic
- Configure class-based policing to rate-limit traffic
- Configure class-based shaping to rate-limit traffic
- Configure class-based shaping on Frame Relay WAN interfaces to rate-limit traffic

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 1

Understanding Traffic Policing and Shaping

Overview

You can use traffic policing to control the maximum rate of traffic sent or received on an interface. Traffic policing is often configured on interfaces at the edge of a network to limit traffic into or out of the network. You can use traffic shaping to control the traffic going out an interface in order to match its flow to the speed of the remote target interface and to ensure that the traffic conforms to policies contracted for it. Traffic policing and traffic shaping differ in the way they respond to traffic violations. Policing typically drops traffic, while shaping typically queues excess traffic. This lesson describes the traffic-policing and traffic-shaping quality of service (QoS) mechanisms that are used to limit the available bandwidth to traffic classes. Because both traffic policing and traffic shaping use the token bucket metering mechanism, this lesson also describes how a token bucket works.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Objectives

Upon completing this lesson, you will be able to explain how to use traffic policing and traffic shaping to condition traffic. This ability includes being able to meet these objectives:

- Describe the purpose of traffic conditioning using traffic policing and traffic shaping
- List key benefits of traffic conditioning using traffic policing and traffic shaping
- Differentiate between the features of traffic policing and traffic shaping
- Explain how a token bucket can be used by network devices to measure traffic rates
- Explain how traffic can be policed using a single token bucket scheme
- Explain how traffic can be policed using a dual token bucket scheme
- Explain how traffic can be policed using a dual-rate metering scheme
- Explain how traffic can be shaped using a single token bucket scheme
- Describe the key traffic policing and shaping mechanisms available in Cisco IOS software and how each compares to the others
- Identify the points in a network where rate-limiting can most effectively be employed

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Traffic Policing and Shaping Overview

This topic describes the purpose of traffic conditioning using traffic policing and traffic shaping.

Both traffic shaping and policing mechanisms are traffic-conditioning mechanisms that are used in a network to control the traffic rate. Both mechanisms use classification so that they can differentiate traffic. They both measure the rate of traffic and compare that rate to the configured traffic-shaping or traffic-policing policy.

The difference between traffic shaping and policing can be described in terms of their implementation:

- Traffic shaping buffers excessive traffic so that the traffic stays within the desired rate. With traffic shaping, traffic bursts are smoothed out by queuing the excess traffic to produce a steadier flow of data. Reducing traffic bursts helps reduce congestion in the network.
- Traffic policing drops excess traffic in order to control traffic flow within specified rate limits. Traffic policing does not introduce any delay to traffic that conforms to traffic policies. Traffic policing can cause more TCP retransmissions, because traffic in excess of specified limits is dropped.

Traffic-policing mechanisms such as class-based policing or committed access rate (CAR) also have marking capabilities in addition to rate-limiting capabilities. Instead of dropping the excess traffic, traffic policing can alternatively mark and then send the excess traffic. This allows the excess traffic to be re-marked with a lower priority before the excess traffic is sent out. Traffic shapers, on the other hand, do not remark traffic; these only delay excess traffic bursts to conform to a specified rate.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Why Use Traffic Conditioners?

This topic describes the key benefits of traffic conditioning using traffic policing and traffic shaping.

Why Use Policing?

- **To limit access to resources when high-speed access is used but not desired (sub-rate access)**
- **To limit the traffic rate of certain applications or traffic classes**
- **To mark down (re-color) exceeding traffic at Layer 2 or Layer 3**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-3

Traffic policing is typically used to satisfy one of these requirements:

- Limiting the access rate on an interface when high-speed physical infrastructure is used in transport. Rate limiting is typically used by service providers to offer customers sub-rate access. For example, a customer may have an OC-3 connection to the service provider but pay only for a T1 access rate. The service provider can rate-limit the customer traffic to T1 speed.
- Engineering bandwidth so that traffic rates of certain applications or classes of traffic follow a specified traffic rate policy. For example, rate-limiting traffic from file-sharing applications to 64 kbps maximum.
- Re-marking excess traffic with a lower priority at Layer 2 and Layer 3, or both, before sending the excess traffic out. Cisco class-based traffic policing can be configured to mark packets at both Layer 2 and Layer 3. For example, excess traffic can be re-marked to a lower differentiated services code point (DSCP) value and also have the Frame Relay discard eligible (DE) bit set before the packet is sent out.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Why Use Shaping?

- **To prevent and manage congestion in ATM and Frame Relay networks, where asymmetric bandwidths are used along the traffic path.**
- **To regulate the sending traffic rate to match the subscribed (committed) rate in Frame Relay or ATM networks.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-4

Traffic shaping is typically used for the following:

- To prevent and manage congestion in ATM and Frame Relay networks, where asymmetric bandwidths are used along the traffic path. If shaping is not used, buffering can occur at the slow (usually the remote) end, which can lead to queuing, causing delays, and overflow, causing drops.
- To prevent dropping of noncompliant traffic by the ATM or Frame Relay service provider by not allowing the traffic to burst above the subscribed (committed) rate. This allows the customer to keep local control of traffic regulation.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Traffic Policing

You can use traffic policing to divide the shared resource (the upstream WAN link) between many flows. In this example, the router FastEthernet interface has an input traffic-policing policy applied to it, in which the mission-critical server traffic rate is not rate-limited, but the User X file-sharing application traffic is rate-limited to 56 kbps. All file-sharing application traffic from User X that exceeds the rate limit of 56 kbps will be dropped.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Traffic Shaping

Traffic-shaping tools limit the transmit rate from a source by queuing the excess traffic. This limit is typically a value lower than the line rate of the transmitting interface. Traffic shaping can be used to account for speed mismatches that are common in nonbroadcast multiaccess (NBMA) networks such as Frame Relay and ATM.

In the figure, these two types of speed mismatches are shown:

- The central site can have a higher speed link than the remote site. You can deploy traffic shaping at the central-site router to shape the traffic rate out of the central-site router to match the link speed of the remote site. For example, the central router can shape the permanent virtual circuit (PVC) outgoing traffic rate (going to the top remote-site router) to 128 kbps to match that remote-site link speed. At each remote-site router, traffic shaping is also implemented to shape the remote-site outgoing traffic rate to 128 kbps to match the committed information rate (CIR).
- The aggregate link speed of all the remote sites can be higher than the central-site link speed, thereby over-subscribing the central-site link speed. In this case, you can configure the remote-site routers for traffic shaping to avoid oversubscription at the central site. For example, you can configure the bottom two remote-site routers to shape the PVC outgoing traffic rate to 256 kbps to avoid the central-site router from being over-subscribed.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Policing vs. Shaping

This topic describes the difference between the features of traffic policing and traffic shaping.

Shaping queues excess traffic by holding packets inside a shaping queue. Use traffic shaping to shape the outbound traffic flow when the outbound traffic rate is higher than a configured shape rate. Traffic shaping smoothes traffic by storing traffic above the configured rate in a shaping queue. Therefore, shaping increases buffer utilization on a router and causes unpredictable packet delays. Traffic shaping can also interact with a Frame Relay network, adapting to indications of Layer 2 congestion in the WAN. For example, if the backward explicit congestion notification (BECN) bit is received, the router can lower the rate limit to help reduce congestion in the Frame Relay network.

You can apply policing to either the inbound or outbound direction, while you can only apply shaping in the outbound direction. Policing drops nonconforming traffic instead of queuing the traffic like shaping. Policing also supports marking of traffic. Traffic policing is more efficient in terms of memory utilization than traffic shaping because no additional queuing of packets is needed.

Both traffic policing and traffic shaping ensure that traffic does not exceed a bandwidth limit, but each mechanism has different impacts on the traffic:

- Policing drops packets more often, generally causing more retransmissions of connection-oriented protocols such as TCP.
- Shaping adds variable delay to traffic, possibly causing jitter.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Measuring Traffic Rates

This topic describes how a token bucket can be used by network devices to measure traffic rates.

The token bucket is a mathematical model that is used by routers and switches to regulate traffic flow. The model has two basic components:

- **Tokens:** Each token represents permission to send a fixed number of bits into the network. Tokens are put into a token bucket at a certain rate by the Cisco IOS software.
- **Token bucket:** A token bucket has the capacity to hold a specified number of tokens. Each incoming packet, if forwarded, takes tokens from the bucket, representing the packet size. If the bucket fills to capacity, newly arriving tokens are discarded. Discarded tokens are not available to future packets. If there are not enough tokens in the token bucket to send the packet, the traffic conditioning mechanisms may take these actions:
  - Wait for enough tokens to accumulate in the bucket (traffic shaping)
  - Discard the packet (traffic policing)

Using a single token bucket model, the measured traffic rate can conform or exceed the specified traffic rate. The measured traffic rate is conforming if there are enough tokens in the single token bucket to transmit the traffic. The measured traffic rate is exceeding if there are not enough tokens in the single token bucket to transmit the traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The figure shows a single token bucket traffic-policing implementation. Starting with a current capacity of 700 bytes worth of tokens accumulated in the token bucket, when a 500-byte packet arrives at the interface, its size is compared to the token bucket capacity (in bytes). The 500-byte packet conforms to the rate limit (500 bytes < 700 bytes), and the packet is forwarded—500 bytes worth of tokens are taken out of the token bucket, leaving 200 bytes worth of tokens for the next packet.

When the next 300-byte packet arrives immediately after the first packet, and no new tokens have been added to the bucket (which is done periodically), the packet exceeds the rate limit. The current packet size (300 bytes) is greater than the current capacity of the token bucket (200 bytes), and the exceed action is performed. The exceed action can be to drop or mark the packet when traffic policing.

Example: Token Bucket as a Piggy Bank

Think of a token bucket as a piggy bank. Every day you can insert a dollar into the piggy bank (the token bucket). At any given time, you can only spend what you have saved up in the piggy bank. On the average, if your saving rate is a dollar per day, your long-term average spending rate will be one dollar per day if you constantly spend what you saved. However, if you do not spend any money on a given day, you can build up your savings in the piggy bank up to the maximum limit that the piggy bank can hold. For example, if the size of the piggy bank is limited to hold five dollars and if you save and do not spend for five straight days, the piggy bank will contain five dollars. When the piggy bank fills to its capacity, you will not be able to put any more money in it. Then, at any time, you can spend up to five dollars (bursting above the long-term average rate of one dollar per day).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Conforming rate, using the piggy bank example, means that if you have two dollars in the piggy bank and you try to spend one dollar, that is considered conforming because you are not spending more than what you have saved in the piggy bank.

Exceeding rate, using the piggy bank example, means that if you have two dollars in the piggy bank and you try to spend three dollars, it is considered exceeding because you are spending more than what you have saved in the piggy bank.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Single Token Bucket Class-Based Policing

This topic describes how traffic can be policed using a single token bucket scheme.

Token bucket operations rely on parameters such as CIR, committed burst (Bc), and committed time window (Tc). Bc is known as the normal burst size. The mathematical relationship between CIR, Bc, and Tc is as follows:

```
CIR (bps) = Bc (bits) / Tc (sec)
```

With traffic policing, new tokens are added into the token bucket based on the interpacket arrival rate and the CIR. Every time a packet is policed, new tokens are added back into the token bucket. The number of tokens added back into the token bucket is calculated as follows:

```
(Current Packet Arrival Time – Previous Packet Arrival Time) *
CIR
```

An amount (Bc) of tokens is forwarded without constraint in every time interval (Tc). For example, if 8000 bits (Bc) worth of tokens are placed in the bucket every 250 milliseconds (Tc), the router can steadily transmit 8000 bits every 250 milliseconds if traffic constantly arrives at the router.

```
CIR (normal burst rate) = 8000 bits (Bc) / 0.25 seconds (Tc) =
32 kbps
```

Without any excess bursting capability, if the token bucket fills to capacity (Bc of tokens), the token bucket will overflow and newly arriving tokens will be discarded. Using the example, in which the CIR is 32 kbps (Bc = 8000 bits and Tc = 0.25 seconds), the maximum traffic rate can never exceed a hard rate limit of 32 kbps.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Dual Token Bucket Class-Based Policing

This topic describes how traffic can be policed using a dual token bucket scheme.

You can configure class-based traffic policing to support excess bursting capability. With excess bursting, after the first token bucket is filled to Bc, extra (excess) tokens can be accumulated in a second token bucket. Excess burst (Be) is the maximum amount of excess traffic over and above Bc that can be sent during the time interval *after* a period of inactivity. With a single rate-metering mechanism, the second token bucket with a maximum size of Be fills at the same rate (CIR) as the first token bucket. If the second token bucket fills up to capacity, no more tokens can be accumulated and the excess tokens are discarded.

When using a dual token bucket model, the measured traffic rate can be as follows:

- **Conforming:** There are enough tokens in the first token bucket with a maximum size of Bc.
- **Exceeding:** There are not enough tokens in the first token bucket but there are enough tokens in the second token bucket with a maximum size of Be.
- **Violating:** There are not enough tokens in the first or second token bucket.

With dual token bucket traffic policing, the typical actions performed are sending all conforming traffic, re-marking (to a lower priority), sending all exceeding traffic, and dropping all violating traffic. The main benefit of using a dual token bucket method is the ability to distinguish between traffic that exceeds the Bc but not the Be. This enables a different policy to be applied to packets in the Be category. Referring to the piggy bank example, think of the CIR as the savings rate ($1 per day). Bc is how much you can save into the piggy bank per day ($1). Tc is the interval at which you put money into the piggy bank (one day). Be ($5) allows you to burst over the average spending rate of one dollar per day if you are not spending a dollar per day.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Dual Token Bucket (Single Rate) Class-Based Policing (Cont.)

- **Traffic is conforming, exceeding, or violating**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-12

Using a dual token bucket model allows traffic exceeding the normal burst rate (CIR) to be metered as exceeding and traffic that exceeds the excess burst rate to be metered as violating traffic. Different actions can then be applied to the conforming, exceeding, and violating traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Dual-Rate Token Bucket Class-Based Policing

This topic describes how traffic can be policed using a dual-rate metering scheme.

With dual-rate metering, traffic rate can be enforced according to two separate rates: CIR and peak information rate (PIR). Before this feature was available, you could meter traffic using a single rate based on the CIR with single or dual buckets. Dual-rate metering supports a higher level of bandwidth management and supports a sustained excess rate based on the PIR.

With dual-rate metering, the PIR token bucket fills at a rate based on the packet arrival rate, and the configured PIR and the CIR token bucket fills at a rate based on the packet arrival rate and the configured CIR.

When a packet arrives, the PIR token bucket is first checked to see if there are enough tokens in the PIR token bucket to send the packet. The violating condition occurs if there are not enough tokens in the PIR token bucket to transmit the packet. If there are enough tokens in the PIR token bucket to send the packet, then the CIR token bucket is checked. The exceeding condition occurs if there are enough tokens in the PIR token bucket to transmit the packet but not enough tokens in the CIR token bucket to transmit the packet. The conforming condition occurs if there are enough tokens in the CIR bucket to transmit the packet.

Dual-rate metering is often configured on interfaces at the edge of a network to police the rate of traffic entering or leaving the network. In the most common configurations, traffic that conforms is sent and traffic that exceeds is sent with a decreased priority or is dropped. Users can change these configuration options to suit their network needs.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Dual-Rate Class-Based Policing (Cont.)

Two-rate policer marks packets as either conforming, exceeding, or violating a specified rate.

- **If (B > Tp), the packet is marked as violating the specified rate**
- **If (B > Tc), the packet is marked as exceeding the specified rate, and the Tp token bucket is updated as Tp = Tp – B**
- **If the packet is marked as conforming to the specified rate, and both token buckets (Tc and Tp) are updated as Tp = Tp – B and Tc = Tc – B.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-14

In addition to rate limiting, traffic policing using dual-rate metering allows marking of traffic according to whether the packet conforms, exceeds, or violates a specified rate.

The token bucket algorithm provides users with three different actions for each packet: a conform action, an exceed action, and an optional violate action. Traffic entering the interface with two-rate policing configured is placed into one of these categories. Within these three categories, users can decide packet treatments. For example, a user may configure a policing policy as follows:

- Conforming packets are transmitted. Packets that exceed may be transmitted with a decreased priority; packets that violate are dropped.
- The violating condition occurs if there are not enough tokens in the PIR bucket to transmit the packet.
- The exceeding condition occurs if there are enough tokens in the PIR bucket to transmit the packet but not enough tokens in the CIR bucket to transmit the packet. In this case, the packet can be transmitted and the PIR bucket is updated to Tp – B remaining tokens, where Tp is the size of the PIR bucket and B is the size of the packet to be transmitted.
- The conforming condition occurs if there are enough tokens in the CIR bucket to transmit the packet. In this case, the packets are transmitted and both buckets (Tc and Tp) are decremented to Tp – B and to Tc – B, respectively, where Tc is the size of the CIR bucket, Tp is the size of the PIR bucket, and B is the size of the packet to be transmitted.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Dual-Rate Token Bucket as a Piggy Bank

Using a dual-rate token bucket is like using two piggy banks, each with a different savings rate. However, you can only take out money from one of the piggy banks at a time.

For example, you can save ten dollars per day into the first piggy bank (PIR = peak spending rate = $10 per day) and then at the same time, you can save five dollars per day into the second piggy bank (CIR = normal average spending rate = $5 per day). However, the maximum amount you can spend is $10 per day, not $15 per day, because you can only take out money from one piggy bank at a time.

In this example, after one day of savings, your first piggy bank (PIR bucket) will contain $10 and your second piggy bank (CIR bucket) will contain $5. The three different spending cases are examined here to show how dual-rate metering operates, using the piggy bank example:

- **Case 1:** If you try to spend $11 at once, then you are violating (Tp < B) your peak-spending rate of $10 per day. In this case, you will not be allowed to spend the $11 because $11 is greater than the $10 you have in the first piggy bank (PIR bucket). Remember, you can only take out money from one of the piggy banks at a time.
- **Case 2:** If you try to spend $9 at once, then you are exceeding (Tp > B > Tc) your normal average spending rate of $5 per day. In this case, you will be allowed to spend the $9 and just the first piggy bank (PIR bucket) will be decremented to $10 – $9, or $1.

  After spending $9, the maximum amount you can continue to spend on that day is decremented to $1.
- **Case 3:** If you try to spend $4, then you are conforming (Tp > B and Tc > B) to your normal average spending rate of $5 per day. In this case, you will be allowed to spend the $4, and both piggy banks (PIR and CIR bucket) will be updated.

The first piggy bank (PIR bucket) will be updated to $10 – $4 = $6, and the second piggy bank (CIR bucket) will be updated to $5 – $4 = $1.

Both piggy banks are updated because after spending $4, the maximum amount you can continue to spend on that day is decremented to $6, and the normal spending rate for that same day is decremented to $1.

Therefore, after spending $4, the following will occur:

- If you continue to spend $7 on that same day, then you will be violating your peak-spending rate for that day. In this case, you will not be allowed to spend the $7 because $7 is greater than the $6 you have in the first piggy bank (PIR bucket).
- If you continue to spend $5 on that same day, then you will be exceeding your normal average spending rate for that day. In this case, you will be allowed to spend the $5 and just the first piggy bank (PIR bucket) will be decremented to $6 – $5, or $1.
- If you continue to spend 50 cents on that same day, then you will be conforming to your normal average spending rate for that day. In this case, you will be allowed to spend the 50 cents, and both piggy banks (PIR and CIR bucket) will be updated. The first piggy bank (PIR bucket) will be updated to $6 – $0.50 = $5.50, and the second piggy bank (CIR bucket) will be updated to $1 – $0.50 = $0.50.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based Traffic Shaping

This topic describes how traffic can be shaped using a single token bucket scheme.

Cisco class-based traffic shaping only applies for outbound traffic.

Class-based traffic shaping uses the basic token bucket mechanism, in which Bc of tokens are added at every Tc time interval. The maximum size of the token bucket is Bc + Be. You can think of the traffic shaper operation like opening and closing of a transmit gate at every Tc interval. If the shaper gate is opened, the shaper checks to see if there are enough tokens in the token bucket to send the packet. If there are enough tokens, the packet is immediately forwarded. If there are *not* enough tokens, the packet is queued in the shaping queue until the next Tc interval. If the gate is closed, the packet is queued behind other packets in the shaping queue.

For example, on a 128-kbps link, if the CIR is 96 kbps, the Bc is 12 kbps, the Be is 0, and the Tc = 0.125 seconds, then during each Tc (125 ms) interval, the traffic shaper gate opens and up to 12 KB can be sent. To send 12 KB over a 128-kbps line will only take 91.25 ms. Therefore the router will, on the average, be sending at three-quarters of the line rate (128 kbps * 3/4 = 96 kbps).

Traffic shaping also includes the ability to send more than Bc of traffic in some time intervals after a period of inactivity. This extra number of bits in excess to the Bc is called Be.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Cisco IOS Traffic Policing and Shaping Mechanisms

This topic describes the key traffic policing and shaping mechanisms available in Cisco IOS software and how each compares to the others.

Cisco IOS Traffic-Policing Mechanisms

	Class-Based Policing
Enable Method	Enabled in policy map
Conditions Actions	Conform, exceed, violate Drop, set, transmit
Implementations	Single or dual token bucket Single or dual rate policing Multi actions

0170_d86

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-16

The figure lists the characteristics of the class-based traffic-policing mechanism that is available in Cisco IOS software. Class-based policing is also available on some Cisco Catalyst switches.

Class-based policing supports a single or dual token bucket. Class-based policing also supports single-rate or dual-rate metering and multiaction policing. Multiaction policing allows more than one action to be applied; for example, marking the Frame Relay DE bit and also the DSCP value before sending the exceeding traffic.

Class-based policing is configured using Modular QoS command-line interface (CLI), or MQC, with the **police** command under the policy map.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Cisco IOS Traffic-Shaping Mechanisms

	CB Shaping	DTS	FRTS
Restriction	Shaper for any subinterface	Shaper on VIP	Shaper for FR only
Classification	Class-based	Subinterface or group or class-based	Per DLCI or subinterface
Link Fragmentation and Interleaving	No Support for FRF.12	Supports FRF.12	Supports FRF.12
Frame Relay Support	Understands BECN and FECN	Understands BECN and FECN	Understands BECN and FECN
Configuration	Supported via MQC	Supported via MQC	Supported via MQC

017G_806

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-17

The figure lists some of the different traffic-shaping mechanisms available in Cisco IOS software: the newer class-based traffic shaping, Distributed Traffic Shaping (DTS), and FRTS.

Class-based traffic shaping uses MQC to allow traffic to be shaped per traffic class as defined by the class map. You can use class-based traffic shaping in combination with class-based weighted fair queuing (CBWFQ), in which the shaped rate is used to define an upper rate limit while the bandwidth statement within the CBWFQ configuration is used to define a minimum rate limit.

DTS is a feature that is specific to the higher-end platforms, such as the Cisco 7500 or the Cisco 12000 series routers. These platforms have the ability to offload traffic shaping from the main processor to the individual interface processor (VIP or line card). In networks in which distributed Cisco Express Forwarding (dCEF) is the preferred mode of switching, DTS on the VIP or line card is the logical choice for implementing traffic shaping.

FRTS is used to shape Frame Relay traffic only. FRTS allows an individual PVC (data-link connection identifier [DLCI]) to be shaped. FRTS can use priority queuing (PQ), custom queuing (CQ), or weighted fair queuing (WFQ) as the shaping queue and only supports FIFO as the software queue.

Both DTS and FRTS support FRF.12 Frame Relay fragmentation, while class-based shaping does not support FRF.12 fragmentation for Frame Relay.

All these traffic-shaping mechanisms can interact with a Frame Relay network, adapting to indications of Layer 2 congestion in the WAN. For example, if the BECN bit is received, the router can lower the rate limit to help reduce congestion in the Frame Relay network. And if the forward explicit congestion notification (FECN) bit is received, the router can generate a test frame with the BECN bit set. This enables the sender to notice congestion even if there is no data traffic flowing back from the receiver to the sender.

Only class-based shaping configurations are described in this module.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Applying Traffic Conditioners

This topic describes the points in a network where rate-limiting can most effectively be employed.

In a typical enterprise network, traffic policing is often implemented at the access or distribution layer to limit certain traffic classes before that traffic exits the campus onto the WAN. Traffic shaping is often implemented at the WAN edge when there are speed mismatches or oversubscription.

In a typical service provider network, traffic policing is often implemented inbound at the provider edge (PE) router to rate-limit incoming traffic from the customer edge (CE) router to ensure that the customer traffic rate is not exceeding the contractual rate. Traffic shaping is often implemented outbound at the PE and at the CE to limit the traffic rate between the PE and CE and to allow for FRF.12 fragmentation on Frame Relay connections between the PE and the CE.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- Traffic shaping and policing are mechanisms that use classification to limit traffic rate.
- Traffic shaping queues excess packets to stay within the contractual rate. Traffic policing typically drops excess traffic to stay within the limit; alternatively, it can re-mark, then send excess traffic.
- Both traffic policing and shaping ensure that traffic does not exceed a bandwidth limit, but they have different impacts on the traffic.
- Traffic rate is metered using a token bucket mathematical model.
- With traffic policing, new tokens are added into the token bucket based on the CIR.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-19

Summary (Cont.)

- Class-based traffic policing can be configured to support excess bursting capability. With excess bursting, after the first token bucket is filled to Bc, extra (excess) tokens can be accumulated in a second token bucket.
- With dual-rate metering, traffic rate can be enforced according to two separate rates: CIR and PIR.
- Class-based traffic shaping uses the basic token bucket mechanism, in which Bc of tokens are added at every Tc time interval.
- Class-based policing is the latest Cisco IOS traffic-policing mechanism. Class-based shaping, DTS, and FRTS are three Cisco IOS traffic-shaping mechanisms.
- Policing is often implemented at the access or distribution layer, while shaping is implemented at the WAN edge.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-20

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 2

Configuring Class-Based Policing

Overview

Cisco IOS software supports two different traffic-policing mechanisms: committed access rate (CAR) and class-based policing. CAR is an older Cisco traffic-policing feature, and class-based policing is a newer Cisco traffic-policing mechanism based on the modular quality of service (QoS) command-line interface (CLI), or MQC. Cisco recommends that you use MQC features when possible to implement QoS in the network. You should avoid using traffic-policing configurations with CAR, because no new features or functionality are planned. However, Cisco will continue to support CAR for existing implementations.

This lesson describes the configuration tasks that are used to implement class-based traffic policing to rate-limit certain traffic classes.

Objectives

Upon completing this lesson, you will be able to configure class-based policing to rate-limit traffic. This ability includes being able to meet these objectives:

- Explain the key features of class-based policing
- Identify the Cisco IOS commands required to configure single-rate class-based policing
- Identify the Cisco IOS commands required to configure dual-rate class-based policing
- Identify the Cisco IOS commands required to configure percentage-based class-based policing
- Identify the Cisco IOS commands used to monitor class-based policing

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based Policing Overview

This topic describes the key features of class-based policing.

Class-Based Policing Overview

- **Class-based policing is used to rate-limit a traffic class to a configured bit rate.**
- **Class-based policing can drop or re-mark and transmit exceeding traffic.**
- **Class-based policing can be implemented using a single or dual token bucket scheme.**
- **Class-based policing supports multiaction policing:**
  - **Applying two or more set parameters as a conform or exceed or violate action**
- **Class-based policing conforms to two RFCs:**
  - **RFC 2697, "A Single Rate Three Color Marker"**
  - **RFC 2698, "A Two Rate Three Color Marker"**
- **Class-based policing is configured using the MQC method.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-2

The class-based policing feature performs these functions:

- Limits the input or output transmission rate of a class of traffic based on user-defined criteria
- Marks packets by setting different Layer 2 or Layer 3 markers, or both

You can implement class-based policing using a single or double token bucket method as the metering mechanism. When the violate action option is not specified in the **police** MQC command, the single token bucket algorithm is engaged; when the violate action option is specified in the **police** MQC command, the dual token bucket algorithm is engaged.

A dual token bucket algorithm allows traffic to do the following:

- Conform to the rate limit when the traffic is within the average bit rate
- Exceed the rate limit when the traffic exceeds the average bit rate, but does not exceed the allowed excess burst
- Violate the rate limit when the traffic exceeds both the average rate and the excess bursts

Depending on whether the current packet conforms with, exceeds, or violates the rate limit, one or more actions can be taken by class-based policing, as follows:

- **Transmit:** The packet is transmitted.
- **Drop:** The packet is dropped.
- **Set IP Precedence (or DSCP value) and transmit:** The IP precedence or differentiated services code point (DSCP) bits in the packet header are rewritten. The packet is then

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

transmitted. This action can be used to either color (set precedence) or recolor (modify existing packet precedence) the packet.

- **Set QoS group and transmit:** The QoS group is set and the packet forwarded. Because the QoS group is only significant within the local router (that is, the QoS group is not transmitted outside the router), the QoS group setting is used in later QoS mechanisms, such as class-based weighted fair queuing (CBWFQ), and performed in the same router on an outgoing interface.
- **Set MPLS Experimental bits and transmit:** The Multiprotocol Label Switching (MPLS) Experimental bits are set. The packet is then transmitted. These are usually used to signal QoS parameters in an MPLS cloud.
- **Set Frame Relay DE bit and transmit:** The Frame Relay discard eligible (DE) bit is set in the Layer 2 (Frame Relay) header and the packet is transmitted. This setting can be used to mark excessive or violating traffic (which should be dropped with preference on Layer 2 switches) at the edge of a Frame Relay network.
- **Set ATM CLP bit and transmit:** The ATM cell loss priority (CLP) bit is set in the Layer 2 (ATM) header and the packet is transmitted. This setting can be used to mark excessive or violating traffic (which should be dropped with preference on Layer 2 switches) at the edge of an ATM network.

Multiaction policing is a mechanism that can apply more than one action to a packet; for example, setting the DSCP as well as the CLP bit on the exceeding packets.

Class-based policing also supports single- or dual-rate metering. With the two-rate policer, traffic policing can be enforced according to two separate rates: committed information rate (CIR) and peak information rate (PIR). You can specify the use of these two rates, along with their corresponding values, by using two keywords, **cir** and **pir**, in the **police** command.

Cisco class-based policing mechanisms conform to these two DiffServ RFCs:

- **RFC 2697, *A Single Rate Three Color Marker*:** The Single Rate Three Color Marker meters an IP packet stream and marks its packets to one of three states: conform, exceed, or violate. Marking is based on a CIR and two associated burst sizes, a committed burst (Bc) size and an excess burst (Be) size. A packet is marked conform if it does not exceed the Bc, exceed if it does exceed the Bc but not the Be, and violate otherwise.
- **RFC 2698, *A Two Rate Three Color Marker*:** The Two Rate Three Color Marker meters an IP packet stream and marks its packets to one of three states: conform, exceed, or violate. A packet is marked violate if it exceeds the PIR. Otherwise a packet is marked either exceed or conform, depending on whether it exceeds or does not exceed the CIR. This is useful, for example, for ingress policing of a service where a peak rate needs to be enforced separately from a committed rate.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Single-Rate Class-Based Policing

This topic describes the Cisco IOS commands that are required to configure single-rate class-based policing.

Configuring Single-Rate Class-Based Policing

```
router(config-pmap-c)#
police avg-rate [BC [BE]] [conform-action action]
[exceed-action action] [violate-action action]
```

- *avg-rate*: Traffic rate in bps (8000 to 200,000,000)
- B_C: normal burst sets the size in bytes
  - Default is 1500 bytes, or CIR / 32, whichever is higher
- B_E: Excess burst sets the size in bytes
  - Default is B_C
- Action:
  - transmit (default conform action)
  - drop (default exceed and violate action)
  - set-prec-transmit ip-precedence
  - set-dscp-transmit dscp
  - set-qos-transmit qos-group
  - set-mpls-exp-transmit mple-exp
  - set frde-transmit
  - set-clp-transmit

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-3

The MQC-based **police** command defines policing parameters for a specific traffic class. The ***avg-rate*** parameter defines the policed CIR; Bc and Be define the token bucket sizes in bytes; and the action defines an action for conforming, exceeding, and optionally violating traffic.

If Bc (in bytes) is not specified, it will default to CIR / 32, or 1500 bytes, whichever is higher. When using the formula CIR / 32 to calculate the default Bc (in bytes), Cisco IOS software uses a Tc of 0.25 second, where:

```
Bc (in bytes) = (CIR x Tc) / 8
Bc (in bytes) = (CIR x 0.25 seconds) / 8 = CIR / 32
```

If Be (in bytes) is not specified, it will default to Bc. In a single token bucket case, Cisco IOS software ignores the Be value. This means that excess bursting is disabled.

The Be rate can be specified when a violate action is configured, therefore using a dual token bucket. This allows Be to be explicitly configured instead of using the default value of Be = Bc. Be specifies the size of the second (excess) token bucket.

Dual token bucket policing with the violate action was introduced in Cisco IOS Release 12.1(5)T.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based Policing Example: Single Rate, Single Token Bucket

```
class-map www.123.com
 match source-address mac 000d.dddf.0480
!
class-map www.456.com
 match source-address mac 000d.dddc.ad21
!
policy-map ServerFarm
 class www.123.com
  police 512000 conform-action transmit exceed-action drop
 class www.456.com
  police 256000 conform-action transmit exceed-action drop
!
interface FastEthernet 0/0
 service-policy input ServerFarm
```

01f0_431

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-4

The class-based policing configuration example shows two configured traffic classes based on upstream MAC addresses. Traffic from the particular web server, which is classified by its MAC address, is policed to a fixed bandwidth with no excess burst capability using a single token bucket. Conforming traffic is sent as-is and exceeding traffic is dropped. In this case, the www.123.com web server is policed to a rate of 512 kbps and the www.456.com web server is policed to a rate of 256 kbps.

Because the violate action is not specified, this will use a single token bucket scheme and no excess bursting is allowed.

In this example, the committed burst size (Bc) is not specified, and therefore it will default to the 512,000 / 32 (16,000 bytes) and 256,000 / 32 (8000 bytes), respectively.

The default Bc setting can be examined by showing the policy map, as follows. Notice that the Be is not displayed because no excess bursting is allowed using a single token bucket with class-based policing.

```
router#show policy-map ServerFarm
  Policy Map ServerFarm
    Class www.123.com
     police cir 512000 bc 16000
       conform-action transmit
       exceed-action drop
    Class www.456.com
     police cir 256000 bc 8000
       conform-action transmit
       exceed-action drop
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Single Rate, Dual Token Bucket Class-Based Policing

The class-based policing configuration example shows two configured traffic classes based on upstream MAC addresses.

Traffic from the particular web server, which is classified by its MAC address, is policed to a fixed bandwidth with excess burst capability using a dual token bucket, by configuring a violate action. Conforming traffic will be sent as-is, exceeding traffic will be marked to IP Precedence 3 and transmitted, and all violating traffic will be dropped.

In this example, because the violate action is specified, a dual token bucket scheme with excess bursting will be used. The committed burst size (Bc) is not specified, and therefore it will default to the 512,000 / 32 (16,000 bytes) and 256,000 / 32 (8000 bytes), respectively. The excess burst size (Be) is also not specified, and therefore it will default to Bc.

The default Bc and Be settings can be examined by showing the policy map, as follows:

```
router#show policy-map ServerFarm
  Policy Map ServerFarm
    Class www.123.com
     police cir 512000 bc 16000 be 16000
       conform-action transmit
       exceed-action set-prec-transmit 3
       violate-action drop
    Class www.456.com
     police cir 256000 bc 8000 be 8000
       conform-action transmit
```

 © 2006, Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

```
exceed-action set-prec-transmit 3
violate-action drop
```

Example: Multiaction Class-Based Policing

The class-based policing configuration is an example of a multiaction class-based policing.

In this case, all Kazaa2 traffic is policed to 56 kbps. All conforming traffic will be marked with a DSCP value of 8, and then transmitted. All exceeding traffic will be marked with a DSCP value of 0, and the CLP bit in the ATM header will also be set before it is transmitted. All violating traffic will be dropped.

The multiaction feature was introduced in Cisco IOS Release 12.2(8)T and is primarily used so that class-based policing can mark at Layer 2 and Layer 3 at the same time.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Dual-Rate Class-Based Policing

This topic describes the Cisco IOS commands that are required to configure dual-rate class-based policing.

Configuring Dual-Rate Class-Based Policing

```
router(config-pmap-c)#
police {cir cir} [bc conform-burst] {pir pir} [be
peak-burst] [conform-action action] [exceed-action
action] [violate-action action]
```

- **Specifies both the CIR and the PIR for two-rate traffic policing**
- **CIR = committed information rate (bps)**
- **PIR = peak information rate (bps)**
- **The Bc and Be keywords and their associated arguments (conform-burst and peak-burst, respectively) are optional**
- **Available in Cisco IOS Release 12.2(4)T**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-7

Dual-rate class-based policing was introduced in Cisco IOS 12.2(4)T. With dual-rate policing, traffic policing can be enforced according to two separate rates: CIR and PIR. The use of these two rates can be specified, along with their corresponding values, by using two keywords, **cir** and **pir**, in the **police** command. The Bc and Be keywords and their associated arguments (**conform-burst** and **peak-burst**, respectively) are optional. If Bc is not specified, Bc (in bytes) will default to CIR / 32, or 1500 bytes, whichever is higher. If Be is not specified, Be (in bytes) will default to PIR / 32, or 1500 bytes, whichever is higher.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Dual-Rate Class-Based Policing

In the example, dual-rate policing is configured on the HTTP traffic class to limit HTTP traffic to an average committed rate of 500 kbps and a peak rate of 1 Mbps.

Conforming traffic will be sent as-is, traffic exceeding 500 kbps (but not exceeding 1 Mbps) will be marked with IP Precedence 3 and then sent, and all traffic exceeding 1 Mbps will be dropped.

Because the Bc and Be are not specified, the default Bc in bytes will be CIR / 32 (Tc = 0.25 seconds), or 1500 bytes, whichever is higher. The default Be in bytes will be PIR / 32 (Tc = 0.25 seconds), or 1500 bytes, whichever is higher.

In this example, the default Bc = 500,000 / 32 = 15,625 bytes, and the default Be = 1,000,000 / 32 = 31,250 bytes.

The **show policy-map** command can be used to display the default settings of the Bc and Be parameters, as follows:

```
Router #show policy-map police-http
  Policy Map police-http
    Class silver
     police cir 500000 bc 15625 pir 1000000 be 31250
       conform-action transmit
       exceed-action set-prec-transmit 3
       violate-action drop
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Percentage-Based Class-Based Policing

This topic describes the Cisco IOS commands that are required to configure percentage-based class-based policing.

Configuring Percentage-Based Class-Based Policing

```
router(config-pmap-c)#
police cir percent percent [bc conform-burst-in-msec] [pir
percent percent] [be peak-burst-in-msec] [conform-action
action] [exceed-action action] [violate-action action]
```

- **Enables the same policy map for multiple interfaces with different bandwidths**
- **Available in Cisco IOS Release 12.2(13)T**

```
policy-map police1
 class bulk-ftp
  police cir percent 20 pir percent 40
  conform-action set-dscp-transmit af11 exceed-action set-dscp-transmit 0
  violate-action drop
!
interface Ethernet 0/0
 service-policy input police1
!
interface Serial 0/0
 service-policy input police1
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-9

The percentage-based policing feature was introduced in Cisco IOS Release 12.2(13)T. Before this feature, traffic policing was configured on the basis of a user-specified amount of bandwidth available on the interface. Policy maps were then configured on the basis of that specific amount of bandwidth, meaning that separate policy maps were required for each interface with a different bandwidth. The percentage-based policing feature provides the ability to configure traffic policing based on a percentage of bandwidth available on the interface. Configuring traffic policing and traffic shaping in this manner enables customers to use the same policy map for multiple interfaces with differing amounts of bandwidth.

Configuring traffic policing based on a percentage of bandwidth is accomplished by using the **police** command.

The percentage-based policing feature also provides the option of specifying burst sizes in ms instead of in bytes when configuring traffic policing based on a percentage of bandwidth.

In the dual-rate percentage-based policing example, the bulk-ftp traffic class is policed to a CIR of 20 percent and a PIR of 40 percent of the interface bandwidth.

Since Bc and Be are not specified, the default Bc in bytes will be CIR / 32, or 1500 bytes, whichever is higher. The default Be in bytes will be PIR / 32, or 1500 bytes, whichever is higher.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Configuring Percentage-Based Class-Based Policing

In the example, the CIR is 20 percent and the PIR is 40 percent (with no Bc and Be specified) and the policy is applied to an Ethernet (10 Mbps) interface. The CIR and PIR in bps for the Ethernet interface are computed as follows:

```
CIR = 10 Mbps * 0.20 = 2 Mbps
PIR = 10 Mbps * 0.40 = 4 Mbps
```

The default values of the Bc and Be parameters will automatically be set to these settings:

```
Bc = 2 Mbps/32 = 62,500 bytes
Be = 4 Mbps/32 = 125,000 bytes
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring Class-Based Policing

This topic describes the Cisco IOS commands that are used to monitor class-based policing.

Monitoring Class-Based Policing

```
router#show policy interface Ethernet 0/0

Ethernet0/0
  Service-policy input: police1
    Class-map: bulk-ftp (match-all)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: protocol ftp
         police:
           cir 20 %
           cir 2000000 bps, bc 62500 bytes
           pir 40 %
           pir 4000000 bps, be 125000 bytes
         conformed 0 packets, 0 bytes; actions:
           set-dscp-transmit cs1
         exceeded 0 packets, 0 bytes; actions:
           set-dscp-transmit default
         violated 0 packets, 0 bytes; actions:
           drop
         conformed 0 bps, exceed 0 bps, violate 0 bps
...
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-10

Use the **show policy-map interface** command to examine the policy map that is being applied to an interface. The output displays the percentage-based policing policy map example shown in the slide.

The output displays the actual value, calculated by Cisco IOS software, of CIR and PIR in bps and the Bc and Be values in bytes, based on the interface bandwidth and the percentage configurations.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Class-based policing is configured using the MQC method and is used to rate-limit a traffic class to a configured bit rate.**
- **Class-based policing can use a single or dual token bucket metering scheme. If a violate action is configured, a dual token bucket will be used.**
- **Class-based policing supports single- or dual-rate metering. Dual-rate metering allows metering of traffic based on two rates (the PIR and the CIR).**
- **Class-based policing supports single actions or multiactions. Multiaction allows the marking of Layer 2 and Layer 3 information at the same time.**
- **Cisco IOS software can automatically calculate an optimal value of Bc and Be based on the configured policed rate.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-11

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006, Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 3

Configuring Class-Based Shaping

Overview

Traffic shaping allows you to control outgoing traffic on an interface to match the speed of transmission to the speed of the remote interface, and to ensure that the traffic conforms to administrative QoS policies. You can shape traffic adhering to a particular profile to meet downstream requirements, thereby eliminating bottlenecks due to data-rate mismatches.

This lesson describes the tasks that are used to configure class-based traffic shaping in order to rate-limit certain traffic classes.

Objectives

Upon completing this lesson, you will be able to configure class-based shaping to rate-limit traffic. This ability includes being able to meet these objectives:

- Explain the key features of class-based shaping
- Explain how the two rate limits—average rate and peak rate—can be used to rate-limit traffic
- Identify the Cisco IOS commands required to configure class-based shaping
- Identify the Cisco IOS commands used to monitor class-based shaping

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based Shaping Overview

This topic describes the key features of class-based shaping.

Traffic shaping allows you to control the traffic going out from an interface in order to match its transmission speed to the speed of the remote, target interface, and to ensure that the traffic conforms to policies contracted for it. You can shape traffic adhering to a particular profile to meet downstream requirements, thereby eliminating bottlenecks in topologies with traffic-rate mismatches or oversubscriptions. Class-based shaping has these properties:

- Class-based shaping, a version of Generic Traffic Shaping (GTS), is configured via MQC.
- Class-based shaping has no packet-marking capabilities.
- Class-based shaping works by queuing exceeding packets until the packets conform to the configured shaped rate.
- Class-based shaping can also be used within the class-based weighted fair queuing (CBWFQ) queuing system. When operating within CBWFQ, the shaped rate provides an upper bandwidth limit while the bandwidth statement within CBWFQ provides a minimum bandwidth guarantee.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Traffic Shaping Methods

This topic describes how the two rate limits—average rate and peak rate—can be used to rate-limit traffic.

Traffic Shaping Methods

- **Class-based shaping has two shaping methods:**
  - **Shaping to the configured average rate**
  - **Shaping to the peak rate**
- **Shaping to the Average rate means forwarding packets at the configured average rate with allowed bursting up to Be when there are extra tokens available. This is the more common method used.**
- **Shaping to the Peak rate means forwarding packets at the peak rate of up to Bc + Be of traffic at every Tc. However, traffic sent above the CIR may be dropped during network congestion. Peak rate shaping is recommended when:**
  - **Network has additional bandwidth available**
  - **Application tolerates occasional packet loss**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-3

Class-based shaping can be configured in these two ways:

- **Shaping to the configured average rate:** Shaping to the average rate forwards up to a committed burst (Bc) of traffic at every committed time window (Tc) interval, with additional bursting capability when enough tokens are accumulated in the bucket. Bc of tokens are added to the token bucket at every Tc time interval. After the token bucket is emptied, additional bursting cannot occur until tokens are allowed to accumulate, which can occur only during periods of silence or when transmit rate is lower than average rate. After a period of low traffic activity, up to Bc + excess burst (Be) of traffic can be sent.
- **Shaping to the peak rate:** Shaping to the peak rate forwards up to Bc + Be of traffic at every Tc time interval. Bc + Be of tokens are added to the token bucket at every Tc time interval. Shaping to the peak rate sends traffic at the peak rate, which is defined as the average rate multiplied by (1 + Be/Bc). Sending packets at the peak rate may result in dropping in the WAN cloud during network congestion. Shaping to the peak rate is only recommended when the network has additional available bandwidth beyond the CIR and applications can tolerate occasional packet drops.

Average rate-shaping is the more common approach being deployed.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Class-Based Shaping

This topic describes the Cisco IOS commands that are required to configure class-based shaping.

Configuring Class-Based Shaping

```
router(config-pmap-c)#
shape {average | peak} average-bit-rate [Bc] [Be]
```

- **It is recommended that you omit the Bc and Be to let Cisco IOS software select the optimal values for these variables.**

```
router(config-pmap-c)#
shape {average | peak} percent [Bc] [Be]
```

- **Specifies the bandwidth percentage. Valid range is a number from 1 to 100.**
- **It is recommended that you omit the Bc and Be to let Cisco IOS software select the optimal values for these variables.**
- **The** shape **(percent) command cannot be configured for use in nested policy maps on 7500 routers or below.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-4

The **shape average** and **shape peak** commands configure average and peak shaping, respectively. The Bc and Be value in bits can be explicitly configured, or Cisco IOS software can automatically calculate their optimal value. It is not recommended that you configure the Bc and Be in order to let the Cisco IOS algorithm determine the best Bc and Be value to use. Class-based traffic shaping uses a single token bucket with a maximum token bucket size of Bc + Be.

The **shape** ***percent*** command is often used in conjunction with the **bandwidth** and **priority** commands. The **bandwidth** and **priority** commands can be used to calculate the total amount of bandwidth available on an entity (for example, a physical interface). When the **bandwidth** and **priority** commands calculate the total amount of bandwidth available on an entity, the total bandwidth is the bandwidth on the physical interface.

The **shape max-buffers** command (not shown) specifies the maximum size of the shaping queue. The shaping queue will queue excess packets until the packets conform to the shaping rate. If the shaping queue is full, packets are tail-dropped. The default max-buffer size is 1000 packets and is usually not changed.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Average Rate, Peak Rate

The figure shows an example configuration for standalone class-based shaping (no CBWFQ). Citrix traffic is classified into the Shape class.

The Shape class is then shaped to different rates on two interfaces:

- On the Serial 0/0 interface, traffic is shaped to the average rate. The Bc and Be values are not configured, allowing Cisco IOS software to automatically calculate their optimal values.

  The Citrix traffic is shaped to the average rate of 16,000 bps. The resulting automatically determined Be and Bc values will both be 8000 bits with a Tc of 500 ms.

- On the Serial 0/1 interface, the Citrix traffic is shaped to the peak rate. Because the Be and Bc are not specified, Cisco IOS software automatically calculates the optimal value for Tc, Bc, and Be. The shape statement is **shape peak 16000.** The resulting automatically determined Be and Bc values will each be 8000 bits with a Tc of 500 ms. Therefore, the peak rate will be as follows:

  ```
  peak rate   = ave rate * (1 + Be / Bc) = 16000 * (1 + 8000 /
  8000)
  = 16000 * 2 = 32000 bps
  ```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Showing the policy map on the interface will display the average rate, peak rate, Bc, Be, and Tc values:

```
router#show policy-map interface s0/1
 Serial0/0

  Service-policy output: ShapePeak

    Class-map: Shape (match-all)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: protocol citrix
      Traffic Shaping
           Target/Average  Byte   Sustain   Excess    Interval  Increment
             Rate          Limit  bits/int  bits/int  (ms)      (bytes)
            32000/16000     2000   8000      8000      500       2000

        Adapt   Queue   Packets   Bytes     Packets   Bytes     Shaping
        Active  Depth                       Delayed   Delayed   Active
        -          0    0         0         0         0         no

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

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Class-Based Shaping with CBWFQ

Class-Based Shaping Example: Class-Based Shaping with CBWFQ

- Bandwidth **statement provides a minimum bandwidth guarantee for the traffic class.**
- Shape **statement provides a maximum traffic rate for the traffic class, and exceed traffic is queued.**

```
policy-map shape-cbwfq
!
class cust1
 shape average 384000
 bandwidth 256
!
interface Serial 0/0
  service-policy output shape-cbwfq
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-6

Class-based shaping can be used in combination with CBWFQ. The shape rate provides a maximum rate limit for the traffic class, while the bandwidth statement within CBWFQ provides a minimum bandwidth guarantee.

In the example, the cust1 traffic class is shaped to an average rate of 384 kbps, and is guaranteed a minimum of 256 kbps by the CBWFQ **bandwidth** statement. Because the Bc and Be are not specified, they will be automatically calculated by the Cisco IOS software.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Class-Based Shaping Hierarchical Policy Maps

The example uses hierarchical policy maps and configures CBWFQ inside the class-based shaping.

The parent policy is the shape-all policy. This parent policy references a child policy named child-cbwfq.

The parent policy map shape-all specifies an average shape rate of 384 kbps for all the traffic (matched by class-default) and assigns the service policy called child-cbwfq as the child policy.

The shaped traffic is further classified into three distinct traffic subclasses:

- **subclass-x:** The subclass-x class is configured to have a minimum guarantee equal to 50 percent of the shaped bandwidth.
- **subclass-y:** The subclass-y class is configured to have a minimum guarantee equal to 20 percent of the shaped bandwidth.
- **subclass-z:** The subclass-z class is configured to have a minimum guarantee equal to 10 percent of the shaped bandwidth.

For example, in a service provider implementation, all traffic going to a specific customer edge router on the S0/0 interface is shaped, then further classified as follows:

- subclass-x class for interactive traffic
- subclass-y class for business-critical traffic
- subclass-z class for bulk traffic

The traffic going out to that specific customer edge router over the S0/0 interface is rate-limited to 384 kbps total. The interactive traffic has a minimal guarantee of 192 kbps (384 x 0.5), the business-critical traffic has a minimal guarantee of 76.8 kbps (384 x 0.2), and the bulk traffic has a minimal guarantee of 38.4 kbps (384 x 0.1).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring Class-Based Shaping

This topic describes the Cisco IOS commands that are used to monitor class-based shaping.

Monitoring Class-Based Shaping

```
router>
show policy-map policy-map-name
```

- **Displays information about the indicated policy map, including the configuration of all classes for a specified service policy map**

```
router>show policy-map shape-cbwfq

Policy Map shape-cbwfq
    Class cust1
      Traffic Shaping
         Average Rate Traffic Shaping
                 CIR 384000 (bps) Max. Buffers Limit 1000 (Packets)
      Bandwidth 256 (kbps) Max Threshold 64 (packets)
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-8

The **show policy-map** command displays the policy-map configuration.

The output represents the CBWFQ in conjunction with the class-based shaping configuration example shown earlier, in which the cust1 traffic class is shaped to an average rate of 384 kbps, with a default buffer limit of 1000 packets and a minimum bandwidth guarantee of 256 kbps.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring Class-Based Shaping (Cont.)

```
router#show policy-map interface s0/0

Serial0/0
  Service-policy output: shape-cbwfq
     Class-map: cust1 (match-all)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: access-group 101
      Traffic Shaping
           Target/Average   Byte   Sustain   Excess    Interval  Increment
             Rate           Limit  bits/int  bits/int  (ms)      (bytes)
            384000/384000    1968   7872      7872      207       984
        Adapt  Queue     Packets   Bytes     Packets   Bytes     Shaping
        Active Depth                         Delayed   Delayed   Active
        -      0         0         0         0         0         no
      Queueing
        Output Queue: Conversation 266
        Bandwidth 256 (kbps) Max Threshold 64 (packets)
        (pkts matched/bytes matched) 0/0
        (depth/total drops/no-buffer drops) 0/0/0
    Class-map: class-default (match-any)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: any
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-9

The **show policy-map interface** command displays all service policies that are applied to the interface.

The output represents the CBWFQ in conjunction with the class-based shaping configuration example shown earlier, where the cust1 traffic class is shaped to an average rate of 384 kbps and has a minimum bandwidth guarantee of 256 kbps.

Among the settings shown are the class-based shaping parameters like Be, Bc, and Tc, and other statistics.

In the example, the shape average rate is 384 kbps and the automatically calculated values for Bc and Be are 7872 bits each, with a Tc of 207 ms.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Class-based shaping delays exceeding packets rather than dropping them and has no marking capabilities.**
- **Class-based shaping can shape to the configured average rate or to the peak rate for packets satisfying the match criteria for a class. Peak Rate = Average Rate * (1 + Be / Bc).**
- **Class-based shaping is configured using the MQC and various** shape **commands.**
- **The** show policy-map interface **command displays the class-based traffic-shaping statistics.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-10

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 4

Configuring Class-Based Shaping on Frame Relay Interfaces

Overview

Frame Relay traffic shaping (FRTS) builds upon existing Frame Relay support of forward explicit congestion notification (FECN) and backward explicit congestion notification (BECN) on Cisco routers by adding new capabilities that improve the scalability and performance of a Frame Relay network. FRTS can help eliminate bottlenecks in Frame Relay networks that consist of high-speed connections at a central hub site and low-speed connections at branch spoke sites. Using FRTS, you can configure rate enforcement to either the committed information rate (CIR) or some other defined value such as the excess information rate on a permanent virtual circuit (PVC) basis. This lesson describes the tasks that are used to configure class-based traffic shaping in order to rate-limit certain traffic classes on Frame Relay interfaces.

Objectives

Upon completing this lesson, you will be able to configure class-based shaping on Frame Relay WAN interfaces to rate-limit traffic. This ability includes being able to meet these objectives:

- Explain the purpose of the Frame Relay FECN, BECN, and the DE bit
- Explain the use of FECN and BECN as a Frame Relay congestion control mechanism
- Explain how class-based shaping can adapt dynamically to available Frame Relay bandwidth by integrating BECN signals
- Explain the FECN-to-BECN propagation mechanism
- Identify the Cisco IOS commands required to configure Frame Relay adaptive class-based shaping on Frame Relay interfaces
- Identify the Cisco IOS commands used to monitor class-based shaping on Frame Relay interfaces

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Frame Relay Refresher

This topic describes the purpose of the Frame Relay FECN, BECN, and the DE bit.

Frame Relay Refresher

- **Frame Relay explicit congestion notification**
  - **FECN**
  - **BECN**
  - **Set by the Frame Relay switch to notify an end device of congestion in the network**
- **Implicit congestion notification**
  - **Provides to a signal the Frame Relay network to determine which frames to discard in preference to their frames**
  - **DE bit**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-2

When the Frame Relay network becomes congested to the point that it cannot process new data transmissions, it begins to discard frames. These discarded frames are detected and then retransmitted by higher-layer protocols, thus causing more congestion. In an effort to prevent this situation, several mechanisms have been developed to notify user devices at the onset of congestion, so that the offered load may be reduced.

Frame Relay performs congestion notification to its Layer 2 endpoints (Frame Relay DTE) by including congestion signaling inside the Layer 2 frame headers. The FECN, BECN, and discard eligible (DE) bits in the Q.922 header of the frame provide in-band congestion signaling.

The FECN and BECN bits are set by a Frame Relay switch to notify an end device (Frame Relay DTE, which may be a router) that the end device should initiate congestion avoidance procedures.

When there is congestion, the network must decide which frames to discard. The DE bit provides the network with a signal to help determine which frames to discard. The DE bit indicates that a frame may be discarded in preference to other frames, if congestion occurs, to maintain the committed QoS within the network. The DE bit may be set by the user on lower-priority frames. Alternatively, the network may set the DE bit to indicate to other nodes that a frame should be preferentially selected for discard, if necessary.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Frame Relay Congestion Control

This topic describes the use of FECN and BECN as Frame Relay congestion control mechanisms.

A Frame Relay switch can explicitly report congestion in two directions: forward and backward.

When a frame queue inside a switch is congested, the switch will generate congestion signals based on the FECN and BECN bits. If congestion occurs in a queue transmitting toward the main receiver of traffic, FECN signals are sent to the receiving Layer 2 endpoint and BECN signals are sent to the sending Layer 2 endpoint.

FECN and BECN bits are not sent as separate frames, but use bits in the header of data frames.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Frame Relay Congestion Adaptation

This topic describes how class-based shaping can adapt dynamically to available Frame Relay bandwidth by integrating BECN signals.

Frame Relay Congestion Adaptation

- **Class-based shaping can adapt dynamically to available Frame Relay bandwidth by integrating BECN signals:**
  - **The bit rate is reduced when BECN packets are received to reduce the data flow through a congested Frame Relay network.**
  - **The bit rate is gradually increased when the congestion is no longer present (BECN packets are no longer received).**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-4

Class-based shaping is able to respond to Layer 2 congestion in the Frame Relay network by reducing its shaping rate to three-quarters of the current rate, until the Layer 2 Frame Relay network recovers from congestion. When BECN flags are no longer received, the rate is slowly increased to the original shaping rate.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

FECN-to-BECN Propagation

This topic describes the FECN-to-BECN propagation mechanism.

Another adaptation method, FECN-to-BECN propagation, configures a router Frame Relay subinterface to reflect received FECN bits as BECN in Q.922 Test Response messages. This enables the sender to notice congestion in the Layer 2 Frame Relay network, even if there is no data traffic flowing from the receiver back to the sender.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Frame Relay Adaptive Class-Based Shaping

This topic describes the Cisco IOS commands that are required to configure Frame Relay adaptive class-based shaping on Frame Relay interfaces.

Configuring Frame Relay Adaptive Class-Based Shaping

```
router(config-pmap-c)#
shape adaptive min-rate
```

- **Adapts the shaping rate when BECN bits are received.**
- **min-rate: Each BECN bit causes the shaping rate to be reduced to three-quarters of the previous rate but not below the *min-rate*.**
- **This command has effect only if used on Frame Relay interfaces.**

```
router(config-pmap-c)#
shape fecn-adapt
```

- **Responds to FECN bits by creating test frames in the opposite direction with the BECN bit set.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-6

The **shape adaptive** command configures the class-based shaping system to adapt the shaping rate to BECN indications.

The ***min-rate*** (also called ***min-cir***) parameter specifies the minimum shaping rate that is allowed. Typically, this is set to the CIR guaranteed for the traffic class when there is congestion in the network. Additional information about BECN integration is as follows:

- If a router receives any BECNs during the current time interval, the router decreases the transmit rate by 25 percent. The router will continue to decrement with each BECN (limited to one decrement per time interval) until the traffic rate gets to min-rate, where transmit-rate ceases to decrease any further.
- After the traffic rate has decreased, it takes 16 time intervals of receiving no BECNs to start to increase traffic again. The traffic increases by (Be + Bc) / 16. Thus, traffic takes much longer to get back to CIR than it did to drop to min-rate (similar to slow start in TCP/IP).

The **shape fecn-adapt** command configures the class-based shaping system to respond to FECN-marked frames with BECN test frames.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Class-Based Shaping with Frame Relay Adaptation

MQC-Based Frame Relay Traffic Shaping Example

```
policy-map child-cbwfq
 class interactive
  bandwidth 32
 class bulk
  bandwidth 24
!
policy-map parent-shape
 class class-default
  shape average 192000
  shape adaptive 96000
  service-policy child-cbwfq
```

```
interface serial0/0
encapsulation frame-relay
!
interface serial0/0.1 point-to-point
ip address 192.168.1.1 255.255.255.0
frame-relay interface-dlci 100
class shape
!
map-class frame-relay shape
 service-policy output parent-shape
```

This feature was already available for routers in the Cisco 7500 and above product range and became available on lower end platforms (Cisco 1700 series, 2500 series, 2600 series, 3620 router, 3631 router, 3640 router, 3660 router, 3725 router, 3745 router, 7200 series, and 7400 series) in Release 12.2(13)T.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-7

The example shows a typical MQC-based Frame Relay traffic shaping with CBWFQ configuration using hierarchical policy maps.

This is similar to hierarchical policy maps with class-based shaping. The difference is where the parent policy map with the **shape** command is applied. In class-based shaping, the parent policy map is applied to the main serial interface. With MQC-based Frame Relay traffic shaping, the parent policy map is applied to the Frame Relay PVC via a Frame Relay map-class.

If the MQC-based Frame Relay traffic-shaping feature is used, the shaping rate that was configured in the parent policy map using MQC will be used as the total available bandwidth for all traffic classes under the child policy map, if CBWFQ or low-latency queuing (LLQ) is configured. If both the **shape average** and **shape adaptive** commands are used for traffic shaping, the available bandwidth will be based on the parameters specified by the **shape adaptive** command.

In the example, all traffic out of the Frame-Relay PVC data-link connection identifier (DLCI) 100 (matched by class-default) will be shaped using the parent-shape policy map to an average rate of 192 kbps. When the BECN bit is detected, the shape rate will start throttling down to a min-rate of 96 kbps. The parent-shape policy map is applied to the Frame Relay map class called “shape” in this example.

In traditional (non MQC-based) FRTS configurations, all FRTS parameters such as CIR, Bc, Be, and minimum CIR (minCIR) are all defined within the Frame Relay map class.

After the aggregate traffic is shaped, the parent-shape policy map references the child-cbwfq policy map, where the two traffic classes (class interactive and class bulk) are defined with different minimum bandwidth guarantees using CBWFQ. The class interactive traffic will have a minimum bandwidth guarantee of 32 kbps, and the class bulk traffic will have a minimum bandwidth guarantee of 24 kbps.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring MQC-Based Frame Relay Traffic Shaping

This topic describes the Cisco IOS commands that are used to monitor MQC-based shaping.

Monitoring MQC-Based Frame Relay Traffic Shaping

```
router#show policy-map interface
Serial0/0.1: DLCI 100 -

  Service-policy output: parent-shape (2811)

    Class-map: class-default (match-any) (2837/0)
      16 packets, 1666 bytes
      30 second offered rate 0 bps, drop rate 0 bps
      Match: any  (2841)
      Traffic Shaping
        Target    Byte   Sustain   Excess    Interval  Increment
Adapt
        Rate      Limit  bits/int  bits/int  (ms)      (bytes)
Active
        192000    1968   7872      7872      41        984       -

        Queue     Packets   Bytes     Packets   Bytes     Shaping
        Depth                         Delayed   Delayed   Active
        0         8         833       0         0         no
…..cont
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-8

The **show policy-map interface** command is used to display the parent policy shaping statistics as well as the child policy CBWFQ statistics.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring MQC-Based Frame Relay Traffic Shaping (Cont.)

```
Service-policy : child-cbwfq (2813)

        Class-map: interactive (match-all) (2815/6)
          0 packets, 0 bytes
          30 second offered rate 0 bps, drop rate 0 bps
          Match: protocol telnet (2819)
          Weighted Fair Queuing
            Output Queue: Conversation 41
            Bandwidth 32 (kbps) Max Threshold 64 (packets)
            (pkts matched/bytes matched) 0/0
        (depth/total drops/no-buffer drops) 0/0/0

output omitted
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-9

The figure is the continuation of the previous **show policy-map interface** output.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

MQC-Based Frame Relay Traffic Shaping Example on Multipoint Main Interface

MQC-Based Frame Relay Traffic Shaping Example on Multipoint Main Interface

```
policy-map child-cbwfq
 class interactive
  bandwidth 32
 class bulk
  bandwidth 24
!
policy-map parent-shape
 class class-default
  shape average 192000
  shape adaptive 96000
  service-policy child-cbwfq
```

```
interface Serial0/1
 ip address 10.1.1.1 255.255.255.0
 encapsulation frame-relay
 frame-relay class shape
 frame-relay map ip 10.1.1.2 102 broadcast
 frame-relay map ip 10.1.1.3 103 broadcast
 frame-relay map ip 10.1.1.4 104 broadcast
!
map-class frame-relay shape
 service-policy output parent-shape
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-10

The example shows a typical MQC-based Frame Relay traffic shaping with CBWFQ configuration using hierarchical policy maps being applied to a multipoint Frame Relay main interface.

Use the **show policy-map interface** command to display the parent policy-shaping statistics as well as the child policy CBWFQ statistics.

```
router#show policy-map interface s0/1
 Serial0/1: DLCI 102 -
  Service-policy output: parent-shape
    Class-map: class-default (match-any)
      0 packets, 0 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: any
      Traffic Shaping
           Target/Average   Byte   Sustain   Excess    Interval  Increment
             Rate           Limit  bits/int  bits/int  (ms)      (bytes)
           192000/192000    1968   7872      7872      41        984
        Adapt  Queue     Packets   Bytes     Packets   Bytes     Shaping
        Active Depth                         Delayed   Delayed   Active
        BECN   0         0         0         0         0         no
      Service-policy : child-cbwfq
        Class-map: interactive (match-all)
          0 packets, 0 bytes
          5 minute offered rate 0 bps, drop rate 0 bps
          Match: protocol telnet
          Queueing
            Output Queue: Conversation 25
            Bandwidth 32 (kbps) Max Threshold 64 (packets)
            (pkts matched/bytes matched) 0/0
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

```
        (depth/total drops/no-buffer drops) 0/0/0
        Class-map: bulk (match-all)
          0 packets, 0 bytes
          5 minute offered rate 0 bps, drop rate 0 bps
          Match: protocol ftp
          Queueing
            Output Queue: Conversation 26
            Bandwidth 24 (kbps) Max Threshold 64 (packets)
            (pkts matched/bytes matched) 0/0
        (depth/total drops/no-buffer drops) 0/0/0
        Class-map: class-default (match-any)
          0 packets, 0 bytes
          5 minute offered rate 0 bps, drop rate 0 bps
          Match: any
Serial0/1: DLCI 103 -
 Service-policy output: parent-shape
   Class-map: class-default (match-any)
     0 packets, 0 bytes
     5 minute offered rate 0 bps, drop rate 0 bps
     Match: any
     Traffic Shaping
          Target/Average   Byte   Sustain   Excess    Interval  Increment
            Rate           Limit  bits/int  bits/int  (ms)      (bytes)
          192000/192000    1968   7872      7872      41        984
       Adapt  Queue     Packets   Bytes     Packets   Bytes     Shaping
       Active Depth                         Delayed   Delayed   Active
       BECN   0         0         0         0         0         no
     Service-policy : child-cbwfq
       Class-map: interactive (match-all)
         0 packets, 0 bytes
         5 minute offered rate 0 bps, drop rate 0 bps
         Match: protocol telnet
         Queueing
           Output Queue: Conversation 25
           Bandwidth 32 (kbps) Max Threshold 64 (packets)
           (pkts matched/bytes matched) 0/0
       (depth/total drops/no-buffer drops) 0/0/0
       Class-map: bulk (match-all)
         0 packets, 0 bytes
         5 minute offered rate 0 bps, drop rate 0 bps
         Match: protocol ftp
         Queueing
           Output Queue: Conversation 26
           Bandwidth 24 (kbps) Max Threshold 64 (packets)
           (pkts matched/bytes matched) 0/0
       (depth/total drops/no-buffer drops) 0/0/0
       Class-map: class-default (match-any)
         0 packets, 0 bytes
         5 minute offered rate 0 bps, drop rate 0 bps
         Match: any
Serial0/1: DLCI 104 -
 Service-policy output: parent-shape
   Class-map: class-default (match-any)
     0 packets, 0 bytes
     5 minute offered rate 0 bps, drop rate 0 bps
     Match: any
     Traffic Shaping
          Target/Average   Byte   Sustain   Excess    Interval  Increment
            Rate           Limit  bits/int  bits/int  (ms)      (bytes)
          192000/192000    1968   7872      7872      41        984
       Adapt  Queue     Packets   Bytes     Packets   Bytes     Shaping
       Active Depth                         Delayed   Delayed   Active
       BECN   0         0         0         0         0         no
     Service-policy : child-cbwfq
         0 packets, 0 bytes
         5 minute offered rate 0 bps, drop rate 0 bps
         Match: protocol telnet
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

```
  Queueing
    Output Queue: Conversation 25
    Bandwidth 32 (kbps) Max Threshold 64 (packets)
    (pkts matched/bytes matched) 0/0
(depth/total drops/no-buffer drops) 0/0/0
Class-map: bulk (match-all)
  0 packets, 0 bytes
  5 minute offered rate 0 bps, drop rate 0 bps
  Match: protocol ftp
  Queueing
    Output Queue: Conversation 26
    Bandwidth 24 (kbps) Max Threshold 64 (packets)
    (pkts matched/bytes matched) 0/0
(depth/total drops/no-buffer drops) 0/0/0
Class-map: class-default (match-any)
  0 packets, 0 bytes
  5 minute offered rate 0 bps, drop rate 0 bps
  Match: any
```

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **The FECN, BECN, and DE bits in the Q.922 header of the frame provide in-band congestion control.**
- **A Frame Relay switch detecting congestion on the output queue informs the receiver by setting the FECN bit on forwarded frames; the receiver reacts by setting the BECN bit on frames going toward the source.**
- **Class-based shaping can adapt dynamically to available Frame Relay bandwidth by integrating BECN signals.**
- **FECN-to-BECN propagation enables a router Frame Relay interface to reflect received FECN bits as BECN in Q.922 Test Response messages if no data frames are sent in the backward direction.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-11

Summary (Cont.)

- **The** shape adaptive **command configures the class-based shaping system to adapt the shaping rate based on BECN indications.**
- **The** shape fecn-adapt **command configures FECN-to-BECN propagation.**
- **The** show policy-map interface **command displays the Frame Relay class-based shaping statistics.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-12

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 5

Frame Relay Voice-Adaptive Traffic Shaping and Fragmentation

Overview

Prior to Cisco IOS Release 12.2(15)T, adaptive shaping could only be configured to shape traffic rates sent to a Frame Relay cloud if congestion notifications such as backward explicit congestion notification (BECN) or forward explicit congestion notification (FECN) were received. However, in the case of voice traffic, by the time such congestion notifications were received by the sending shaper, the VoIP packet might already have been delayed or dropped. Furthermore, if fragmentation were required on slow-speed data-link connection identifiers (DLCIs), FRF.12 was enabled permanently, regardless of whether voice traffic was present.

As an alternative to congestion-adaptive traffic-shaping, Frame Relay Voice-Adaptive Traffic Shaping (FR VATS) was introduced in Release 12.2(15)T. FR VATS enables adaptive shaping, a rating down to minimum committed information rate (CIR), or minCIR, and the enabling of FRF.12 fragmentation and interleaving only when voice activity is detected on the Frame Relay DLCI. Voice activity is perceived by the presence of packets in the low-latency queueing (LLQ), or the presence of H.323 signaling packets. When voice activity is no longer present, fragmentation is deactivated (after a specified interval) and the shaper resumes shaping to CIR (not minCIR). By default, FR VATS deactivates 30 seconds after voice activity is no longer detected, but this timeout period can be administratively tuned with the deactivation keyword and value. This lesson describes the features and functionality of FR VATS.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Objectives

Upon completing this lesson, you will be able to configure voice-adaptive traffic shaping and fragmentation on Frame Relay WAN interfaces. This ability includes being able to meet these objectives:

- Describe the Frame Relay voice-adaptive traffic shaping and fragmentation feature
- Describe the benefits from deploying voice-adaptive traffic shaping and fragmentation
- Describe the prerequisites for deploying voice-adaptive traffic shaping and fragmentation
- List the router platforms that support FR-VATS
- Describe how the voice-adaptive traffic shaping and fragmentation feature works over a Frame Relay PVC
- Describe the steps in configuring voice-adaptive traffic shaping and fragmentation
- Using Cisco IOS software CLI commands, monitor voice-adaptive traffic over a Frame Relay WAN

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Frame Relay Voice-Adaptive Traffic Shaping and Fragmentation

This topic describes Frame Relay voice-adaptive traffic shaping and fragmentation.

FR-VATS Fragmentation

- **Engaged during congestion**
- **Mechanism activation and deactivation of traffic shaping and fragmentation**
- **Traffic is shaped to minimum CIR**
- **Fragmentation per FRF.12**
- **When not engaged, data can burst to line speed.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-2

To address the needs of voice traffic in a Frame Relay network in which shaping is desired only when voice is present, but otherwise bursting to port-speed is desired, you can configure the FR-VATS feature. This feature was introduced in Cisco IOS Release 12.2(15)T.

FR-VATS monitors the Frame Relay permanent virtual circuit (PVC), and when voice activity is present, this feature automatically shapes the CIR to the minCIR. If no voice is present, the voice adaptive feature is turned off, allowing the data traffic to burst to the line rate, which is the CIR. The voice adaptive feature determines the presence of voice based on packets entering a low-latency queue. The feature then automatically turns traffic shaping on or off. Similarly, the FR-VATS feature is triggered on Frame Relay Fragmentation (FRF.12) on links less than 768 kbps when voice packets are present to prevent unnecessary serialization delays.

Because FR-VATS automatically detects the presence of voice, there could be brief quality degradation in the first couple of seconds of the first voice call made across a PVC. This occurs while the interfaces in the PVC reconfigure themselves to shape to the minimum CIR and empty out their buffers. FR-VATS is not a predictive algorithm, and the change in behavior is triggered by voice packets flowing across the PVC or the presence of H.323 call-setup signaling.

The FR-VATS deactivation period is tunable and, by default, is set to 30 seconds. If you tune it, you should set the timer so that the feature will not turn off frequently during normal business use (for example, between every two calls). The feature works better on PVCs that always have at least one voice call present during daytime use and relinquish shaping only at night.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Benefits in Deploying FR-VATS

This topic describes the benefits of deploying FR-VATS.

Benefits of Deploying FR-VATS

- **Prevents delay of voice packets during congestion**
- **Maximizes utilization of PVC**
- **Reduces router CPU utilization**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-3

Before the introduction of the FR-VATS feature, Frame Relay adaptive shaping was used to reduce the sending rate when a network was congested. Because the adaptive shaping mechanism was triggered by network congestion, voice traffic might already have been delayed by the time the sending rate was reduced. The Frame Relay voice-adaptive traffic shaping and fragmentation feature helps to ensure voice quality by adjusting the rate of traffic based on the presence of voice on the PVC.

Frame Relay voice-adaptive traffic shaping and fragmentation provides these benefits:

- Prevents delay of voice packets when network congestion occurs by reducing the traffic rate to the minCIR and turning on fragmentation when voice packets are present on a PVC
- Maximizes utilization of the PVC by increasing the traffic rate to the CIR when voice packets are not present
- Reduces CPU utilization by turning off fragmentation when there are no voice packets present

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Prerequisites for Deploying FR-VATS

This topic describes the prerequisites for deploying FR-VATS.

Prerequisites in Deploying FR-VATS

Traffic Shaping:

- **Traffic Shaping and LLQ must be configured using MQC.**

Link Fragmentation:

- **End-to-end fragmentation must be configured in a map or on the interface.**
- **LLQ must be configured.**
- **End-to-end fragmentation must be configured on peer router.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-4

A prerequisite for Frame Relay voice-adaptive traffic shaping is that traffic shaping and LLQ must be configured using the Modular quality of service (QoS) command-line interface (CLI), or MQC.

Prerequisites for Frame Relay voice-adaptive fragmentation are as follows:

- End-to-end fragmentation must be configured in a map class or on the interface.
- Frame Relay traffic shaping or traffic shaping using the MQC must be configured. If end-to-end fragmentation is configured on the interface, traffic shaping must be configured using the MQC.
- LLQ must be configured.
- End-to-end fragmentation must be configured on the peer router. Although the peer router may not see the fragmented packets that are expected from the router doing voice-adaptive fragmentation, the peer router will be able to handle large unfragmented packets in addition to fragmented packets.

Note The feature supports FRF.12 fragmentation only. Neither FRF.11 Annex C nor Cisco proprietary fragmentation is supported.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Supported Platforms

This topic describes the router platforms that support FR-VATS.

Supported Platforms

- **Introduced in Cisco IOS software Release 12.2(15)T**
  - **Cisco 1700**
  - **Cisco 2600, Cisco2600XM and Cisco 2800**
  - **Cisco 3600, Cisco 3700, and Cisco 3800**
  - **Cisco 7200, Cisco 7400, and Cisco 7500**
  - **Not supported on Versatile Interface Processors (VIPs)**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-5

The figure lists the Cisco router platforms that support FR-VATS.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Frame Relay Voice-Adaptive Traffic Shaping and Fragmentation Operation

This topic describes how priority packets and H.323 call setup signaling packets trigger the Frame Relay voice-adaptive traffic shaping and fragmentation feature.

The figure shows how FR-VATS is engaged and what occurs on the interface in order for the FR-VATS feature to be used. After being configured, the FR-VATS feature is automatically triggered, throttling back CIR to minCIR, if packets enter the LLQ or if H.323 call-setup signaling is detected. As packets enter the strict priority queue and the circuit is reshaped, fragmentation is engaged at Layer 2 so that priority packets are interleaved with the fragmented data packets. When the strict priority queue is void of packets, the Frame Relay circuit is reshaped back to its CIR, when data packets can burst to CIR or line speed depending on the service level agreement (SLA). Reshaping to the CIR only occurs when there are no strict priority packets in the queue, and in the absence of H.323 call-setup signaling.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring the Frame Relay Voice-Adaptive Feature

This topic describes how to configure voice adaptation with end-to-end fragmentation on the interface.

Configuring Voice Adaptation with End-to-End Fragmentation on the Interface

```
interface serial0
 encapsulation frame-relay
 frame-relay fragmentation voice-adaptive deactivation 30 (FR-VATS + FRF.12)
 frame-relay interface-dlci 100
  class voice-adaptive-class (Binds voice_adaptive_class to the DLCI)
  frame-relay fragment 80 end-to-end (FRF.12 fragment for 384 kbps PVC on the interface)
!
class-map match-all voice
 match access-group 102
class-map match-all data
 match access-group 101
!
policy-map VOICE-TRAFFIC-SHAPING
 class voice
  priority 10 (Strict Priority)
 class data
  bandwidth 10
!
policy-map TRAFFIC-SHAPE
 class class-default
  shape average 729600 3648 0 (CIR = 95% of PVC [768 kbps], Bc to minCIR / 100, Be = 0)
  shape adaptive 364800 (Sets minimum CIR to 95% of 384 kbps)
  shape fr-voice-adapt deactivation 30 (Enables FR-VATS [default timer])
  service-policy VOICE-TRAFFIC-SHAPING (Embedded LLQ to policy)
!
map-class frame-relay voice-adaptive-class
 service-policy output TRAFFIC-SHAPE (Embeds FR-VATS feature to map-class)
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-7

The figure shows an example of how to configure voice adaptation and end-to-end fragmentation on the interface. Notice that the strict priority queue which is used by FR-VATS is configured within the policy map **VOICE-TRAFFIC-SHAPING** and that policy is embedded within the traffic shaper policy map **TRAFFIC-SHAPE**, where the voice-adaptive feature is enabled. In order to tie the voice-adaptive feature to the interface, the policy map **TRAFFIC-SHAPE** must be embedded into the Frame Relay map-class **voice-adaptive-class**. Data-link Connection Identifier (DLCI) 100 is bound to the voice-adaptive feature through the map-class **voice-adaptive-class .** The end-to-end fragmentation configuration can be seen on the interface, and is enabled by the command **frame-relay fragment 80 end-to-end .** The end-to-end fragmentation configuration requires that the peer router have the same end-to-end fragmentation configuration.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Voice Adaptation with End-to-end Fragmentation Through a Map-Class

```
interface serial0
 encapsulation frame-relay
 frame-relay fragmentation voice-adaptive deactivation 30  (FR-VATS + FRF.12)
 frame-relay interface-dlci 100
  class voice-adaptive-class  (Binds voice_adaptive_class to the DLCI)
!
map-class frame-relay voice-adaptive-class
 frame-relay fragment 80  (FRF.12 fragment for 384 kbps PVC through the map-Class)
 service-policy output TRAFFIC-SHAPE (Embeds FR-VATS Feature to map-class)
!
class-map match-all voice
 match access-group 102
class-map match-all data
 match access-group 101
!
policy-map VOICE-TRAFFIC-SHAPING
 class voice
  priority 10        (Strict Priority)
 class data
  bandwidth 10
!
policy-map TRAFFIC-SHAPE
 class class-default
  shape average 729600 3648 0  (CIR=95% of PVC [768 kbps] Bc to minCIR / 100 Be = 0)
  shape adaptive 364800  (Sets minimum CIR to 95% of 384 kbps)
  shape fr-voice-adapt deactivation 30  (Enables FR-VATS with default timer)
  service-policy VOICE-TRAFFIC-SHAPING  (Embedded LLQ to policy)
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-8

The figure shows an example of how to configure the voice-adaptive feature and end-to-end fragmentation on the interface. The only difference between this figure and the previous figure is that end-to-end fragmentation on the previous figure was configured on the interface, and in this example the fragmentation is enabled by being embedded within the Frame Relay map class.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Traffic Shaping and End-to-End Fragmentation – Caveats

- **Frame Relay voice-adaptive traffic shaping:**
  - **Can coexist with adaptive traffic shaping**
  - **Can be used with or without voice-adaptive fragmentation**
- **Frame Relay voice-adaptive fragmentation:**
  - **Can coexist with adaptive traffic shaping**
  - **Can be used with or without voice-adaptive traffic shaping**
- **Frame Relay voice-adaptive traffic shaping and fragmentation will respond to any packets, voice or data, in the priority queue.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-9

Voice-adaptive traffic shaping and end-to-end fragmentation are designed to work together, but they can each coexist with other traffic-shaping configurations. Voice-adaptive traffic shaping and end-to-end fragmentation supports only FRF.12 fragmentation. These are some caveats relative to voice-adaptive traffic shaping and end-to-end fragmentation:

- **Frame Relay voice-adaptive traffic shaping:** Frame Relay voice-adaptive traffic shaping enables a router to reduce the PVC sending rate to the minCIR when packets (usually voice packets) are detected in the low-latency queuing priority queue, or when H.323 call-setup signaling packets are present. When there are no packets in the priority queue and signaling packets are not present for a configured period of time, the router increases the PVC sending rate from minCIR to CIR to maximize throughput. Frame Relay voice-adaptive traffic shaping can be used at the same time as other types of adaptive traffic shaping. For example, if both Frame Relay voice-adaptive traffic shaping and adaptive shaping based on interface congestion are configured, the router will reduce the sending rate to minCIR if there are packets in the priority queue or the interface queue size exceeds the configured threshold. Frame Relay voice-adaptive traffic shaping can be used in conjunction with or independently of voice-adaptive fragmentation.

Note Although the priority queue is generally used for voice traffic, Frame Relay voice-adaptive traffic shaping will respond to any packets (voice or data) in the priority queue.

- **Frame Relay voice-adaptive end-to-end fragmentation:** Frame Relay voice-adaptive fragmentation enables a router to fragment large data packets when packets (usually voice packets) are detected in the low-latency priority queue, or when H.323 call-setup signaling packets are present. When there are no packets in the priority queue for a configured period of time and signaling packets are not present, fragmentation is stopped. Frame Relay voice-adaptive fragmentation can be used in conjunction with or independent of voice-adaptive traffic shaping.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

To use voice-adaptive fragmentation, you must also have end-to-end fragmentation configured in a map class or on the interface.

Note	Although the priority queue is generally used for voice traffic, Frame Relay voice-adaptive fragmentation will respond to any packets (voice or data) in the priority queue.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring the Frame Relay Voice-Adaptive Feature

This topic describes monitoring the FR-VATS feature using Cisco IOS software commands.

Monitoring FR-VATS

- **show policy map**
- **show frame-relay fragment**
- **show frame-relay pvc**
- **show policy-map interface**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-10

The **show policy map** command gives a static verification of the policy configuration only. No dynamic counters are included in its output. The following is a sample output:

```
WAN-EDGE#show policy TRAFFIC-SHAPE
  Policy Map TRAFFIC-SHAPE
    Class class-default
      Traffic Shaping
         Average Rate Traffic Shaping
           CIR 729600 (bps) Max. Buffers Limit 1000 (Packets)
           Bc 3648               Be 0
           Adapt to 364800 (bps)
           Voice Adapt Deactivation Timer 30 Sec
```

Note The highlighted section in the example shows the configuration specifics related to CIR, minCIR, and the default FR-VATS deactivation timer.

The **show frame-relay fragment** command provides verification of the fragment size, whether static fragmentation or voice-adaptive fragmentation is configured for a DLCI. Additionally, dynamic counters monitor how many frames required fragmentation in either direction. The following is a sample output:

```
WAN-EDGE#show frame-relay fragment 102
 interface      dlci  frag-type    frag-size  in-frag    out-frag dropped-frag
 Serial2/0.12   102   end-to-end     480        5476       2035       0
```


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Note The highlighted section shows, on the DLCI 102, the end-to-end fragmentation type, configured fragmentation size on the DLCI, and the total number of fragments sent by the designated DLCI 102.

The **show frame-relay pvc** command provides detailed traffic counters for Frame Relay DLCIs. Also, this command verifies that adaptive fragmentation has been enabled and is currently active on the DLCI. The following is a sample output:

```
WAN-EDGE#show frame-relay pvc 102
PVC Statistics for interface Serial2/0 (Frame Relay DTE)
DLCI = 102, DLCI USAGE = LOCAL, PVC STATUS = ACTIVE, INTERFACE = Serial2/0.12
 input pkts 97269         output pkts 96892        in bytes 34314039
 out bytes 34622446       dropped pkts 0           in pkts dropped 0
 out pkts dropped 19               out bytes dropped 1596
 late-dropped out pkts 19          late-dropped out bytes 1596
 in FECN pkts 0           in BECN pkts 0           out FECN pkts 0
 out BECN pkts 0          in DE pkts 0             out DE pkts 0
 out bcast pkts 62064     out bcast bytes 5315261
 5 minute input rate 150000 bits/sec, 26 packets/sec
 5 minute output rate 151000 bits/sec, 26 packets/sec
 pvc create time 3d02h, last time pvc status changed 00:13:08
 fragment type end-to-end fragment size 480 adaptive active, time left 28 secs
```

Note The highlighted section indicates that there is currently an active end-to-end session and that there are 28 seconds left before the router throttles the Frame Relay circuit back up to the CIR rate.

The **show policy interface** command is the most useful of the **show** commands, and verifies that FR-VATS is active, due to the presence of voice traffic on the link. The following is a sample output:

```
WAN-EDGE#show policy interface serial 2/0.12
 Serial2/0.12: DLCI 102 -

  Service-policy output: MQC-FR-VATS-768

    Class-map: class-default (match-any)
      35164 packets, 30211929 bytes
      30 second offered rate 234000 bps, drop rate 0 bps
      Match: any
      Traffic Shaping
        Target/Average   Byte   Sustain   Excess    Interval  Increment
          Rate           Limit  bits/int  bits/int    (ms)      (bytes)
          729600/729600   456     3648       0          5         456
        Adapt  Queue     Packets   Bytes     Packets   Bytes     Shaping
        Active Depth                         Delayed   Delayed   Active
        BECN   0         41367     30255560  13977     3645051   yes
        Voice Adaptive Shaping active, time left 29 secs
```

Note The highlighted section shows that the FR-VATS feature is active. Also, the time left for FR-VATS to deactivate is included in the output. The output for this command has been truncated, because it is not directly relevant to this example, but if the total output was included, it would show the nested queuing policies assigned to the interface.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- When engaged the voice-adaptive traffic-shaping feature reshapes the link to minimum CIR.
- When engaged, voice-adaptive fragmentation interleaves voice packets with data packets.
- Both voice-adaptive traffic shaping and fragmentation respond to packets in the priority queue.
- Voice-adaptive traffic shaping can coexist with adaptive traffic shaping.
- Voice-adaptive fragmentation can be configured separately from voice-adaptive traffic shaping.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-11

Summary (Cont.)

- Voice-adaptive traffic shaping requires the configuration of a priority queue.
- End-to-end fragmentation can be configured directly on the interface or through a map-class.
- The features were introduced with Cisco IOS Release Software12.2(15)T

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-12

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Summary

This topic summarizes the key points discussed in this module.

Module Summary

- **Traffic shaping queues excess packets to stay within the shape rate. Traffic policing typically drops excess traffic to stay within the rate limit; alternatively, it can re-mark, then send excess traffic. Both traffic shaping and policing measure traffic rate using a token bucket mathematical model.**
- **Class-based policing features include: Drop or re-mark and transmit exceeding traffic, single or dual token bucket, single- or dual-rate policing, multiaction policing.**
- **Class-based shaping can shape to the average or peak rate for a traffic class.**
- **Class-based shaping can throttle down the traffic rate dynamically when BECN bits are received. Class-based shaping can respond to FECN bits by creating test frames in the opposite direction with the BECN bit set.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—7-1

Rate-limiting is the ability to prevent excess traffic from entering or leaving the network. Rate-limiting is required because of speed mismatches, oversubscriptions, sub-rate access, and to prevent unwanted traffic from causing network congestion.

To measure the traffic rate, a token bucket is used. Parameters that define the operations of a token bucket include the CIR, Bc (normal burst size), Be, and Tc.

The CIR can be calculated by using the following formula:

```
CIR (in bps) = Bc (in bits)/ Tc (in seconds)
```

Both traffic policing and traffic shaping are QoS mechanisms that are used to rate-limit a traffic class. Traffic policing operates by dropping excess traffic while traffic shaping delays excess traffic with the aid of queuing.

In Cisco IOS software, the most current rate-limiting mechanisms are class-based policing and class-based shaping. Both of these rate-limiting mechanisms are configured using the MQC. On Frame Relay interfaces, class-based shaping can be used in conjunction with CBWFQ and supports BECN adaptation and FECN-to-BECN propagation.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

References

For additional information, refer to these resources:

- To learn more about configuring class-based policing, refer to *Configuring Traffic Policing* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt4/qcfpoli.htm
- To learn more about metering in class-based policing, refer to *RFC 2697: A Single Rate Three Color Marker* at this URL: http://www.faqs.org/rfcs/rfc2697.html
- To learn more about metering in class-based policing, refer to *RFC 2698: A Two Rate Three Color Marker* at this URL: http://www.faqs.org/rfcs/rfc2698.html
- To learn more about class-based shaping, refer to *Configuring Class-Based Shaping* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt4/qcfcbshp.htm
- To learn more about DTS, refer to *Configuring Distributed Traffic Shaping* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt4/qcfdts.htm
- To learn more about FRTS, refer to *Configuring Frame Relay and Frame Relay Traffic Shaping* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/qos_c/qcpart4/qcfrts.htm
- To learn more about FRTS, refer to *Frame Relay Traffic Shaping* at this URL: http://www.cisco.com/warp/public/125/21.pdf
- To learn more about class-based shaping, refer to *Configuring Class-Based Shaping* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt4/qcfcbshp.htm
- For information about Frame Relay configuration, refer to *Configuring Frame Relay and Frame Relay Traffic Shaping* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/12cgcr/qos_c/qcpart4/qcfrts.htm
- To learn more about traffic policing and traffic shaping, refer to *Part 4: Policing and Shaping* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123cgcr/qos_vcg.htm#1001018
- For information on other traffic shaping mechanisms, refer to the software configuration documentation for your Cisco IOS software release.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check

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

Q1) Which of the following is a major difference between traffic policing versus traffic shaping? (Source: Understanding Traffic Policing and Traffic Shaping)

A) Traffic policing drops excess traffic while traffic shaping delays excess traffic by queuing it.
B) Traffic policing is only applied in the outbound direction while traffic shaping can be applied to both the inbound and outbound directions.
C) Traffic policing is *not* available on the Catalyst switches such as the 2950, while traffic shaping *is* available on Catalyst switches such as the 2950.
D) Traffic policing requires policing queues to buffer excess traffic, while traffic shaping does not require any queues to buffer excess traffic.

Q2) Which mathematical model is used by traffic policing mechanisms to meter traffic? (Source: Understanding Traffic Policing and Traffic Shaping)

A) token bucket
B) RED
C) FIFO metering
D) Predictor or Stacker

Q3) Which three situations typically require the use of a rate-limiting mechanism? (Choose three.) (Source: Understanding Traffic Policing and Traffic Shaping)

A) Frame Relay connection with mismatch speeds at the two end points
B) hub-and-spoke Frame Relay topology with oversubscription at the hub site
C) service provider providing sub-rate access
D) in the service provider high-speed backbone to avoid congestion

Q4) When configuring single-rate class-based policing, which configuration parameter is used to enable a dual token bucket? (Source: Configuring Class-Based Policing)

A) configuring a “violate” action
B) configuring an “exceed” action
C) configuring the PIR in addition to the CIR
D) configuring Be

Q5) What is the main advantage of using multiaction policing? (Source: Configuring Class-Based Policing)

A) to distinguish between exceeding and violating traffic
B) to distinguish between conforming and exceeding traffic
C) to allow the setting of both Layer 2 and Layer 3 QoS markers at the same time
D) to allow marking of the traffic before transmission

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q6) Which two statements are true when class-based shaping is used in conjunction with CBWFQ? (Choose two.) (Source: Configuring Class-Based Shaping)

A) The **bandwidth** command defines the minimum guaranteed bandwidth for the traffic class.
B) The **bandwidth** command defines the maximum guaranteed bandwidth for the traffic class.
C) The **shape peak** command defines the maximum rate limit for the traffic class.
D) The **shape average** command defines the minimum bandwidth guaranteed for the traffic class.

Q7) Which are the two configuration options when configuring class-based traffic shaping? (Choose two.) (Source: Configuring Class-Based Shaping)

A) shape average
B) shape peak
C) single or dual token bucket
D) single or multiaction traffic shaping
E) single- or dual-rate traffic shaping

Q8) Which three items are features of class-based policing? (Choose three.) (Source: Configuring Class-Based Shaping)

A) single or dual token bucket
B) single or multiaction policing
C) single- or dual-rate policing
D) single or dual FIFO queuing
E) single or dual drop threshold

Q9) What is the min-rate? (Source: Configuring Class-Based Shaping on Frame Relay Interfaces)

A) the minimum bandwidth guaranteed by CBWFQ for the traffic class
B) the minimum bandwidth guaranteed by the LLQ for the traffic class
C) the minimum speed the traffic rate will slow down to when the BECN bits are being received
D) the minimum speed the traffic rate will slow down to when the FECN bits are being received
E) the minimum class-based policing traffic rate for a traffic class

Q10) What is the function of FECN-to-BECN propagation? (Source: Configuring Class-Based Shaping on Frame Relay Interfaces)

A) to allow a receiving router to send back a test frame with the BECN bit set to the sender when the receiver has no data to send back to the sender after receiving the FECN bit
B) to allow a router to slow down the traffic rate to the min-rate when either the FECN or BECN bit is received
C) to allow a router to slow down the traffic rate to the min-rate when the BECN bit is received
D) to allow a router to slow down the traffic rate to the min-rate when the FECN bit is received
E) to allow a receiving router to set the BECN bit on a user dataframe going back to the sender after receiving the FECN bit

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check Answer Key

Q1) A

Q2) A

Q3) A, B, C

Q4) A

Q5) C

Q6) A, C

Q7) A, B

Q8) A, B, C

Q9) C

Q10) A

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module 8

Link Efficiency Mechanisms

Overview

In networks consisting of a large WAN covering many sites, the monthly recurring cost of even the smallest link upgrade can be too high. In some cases, the only option for low-bandwidth remote sites is to upgrade their wide-area circuit. In other cases, a set of quality of service (QoS) techniques can be used that improves the efficiency of these low-speed WAN links, which, in this context, typically refer to links of speeds less than or equal to 768 kbps. Header compression and payload compression mechanisms reduce the size of packets, reducing delay and increasing available bandwidth on a link. Other QoS link-efficiency techniques, such as Link Fragmentation and Interleaving (LFI), allow fragile traffic types such as voice and interactive traffic to be sent either ahead or interleaved with larger, more aggressive flows. These techniques decrease latency and assist in meeting the service-level requirements of delay-sensitive traffic types such as voice.

This module describes the different link efficiency mechanisms that are available in Cisco IOS software to implement header compression, payload compression, and LFI.

Module Objectives

Upon completing this module, you will be able to use Cisco link efficiency mechanisms to improve the bandwidth efficiency of low-speed WAN links. This ability includes being able to meet these objectives:

- Explain how link efficiency mechanisms can be used to improve bandwidth efficiency and reduce delay
- Configure class-based TCP and class-based RTP header compression to improve bandwidth efficiency and reduce delay
- Configure LFI to improve bandwidth efficiency and reduce delay

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 1

Understanding Link Efficiency Mechanisms

Overview

Interactive traffic such as Telnet and VoIP are susceptible to increased latency when network processes using large packets, such as bulk FTP, traverse WAN links. Packet delay is especially significant when FTP packets are queued on slower links (in this context, links with speeds less than or equal to 768 kbps) within the WAN. To solve delay problems on slow bandwidth links, a method for fragmenting larger frames and then queuing the smaller frames between fragments of the larger frames is required. To meet this requirement, Cisco IOS software supports PPP Multilink Link Fragmentation and Interleaving (MLP LFI) and Frame Relay Fragmentation (FRF.12). In addition, you can use other tools, such as header and payload compression techniques, to reduce the size of frames that are sent on WAN links.

This lesson describes different approaches for improving the efficiency of WAN links. The lesson describes link efficiency mechanisms that either compress the payload or reduce packet headers. It also describes the different Layer 2 LFI mechanisms and Frame Relay fragmentation.

Objectives

Upon completing this lesson, you will be able to explain how link efficiency mechanisms can be used to improve bandwidth efficiency and reduce delay. This ability includes being able to meet these objectives:

- Explain the various link efficiency mechanisms and their functions
- Describe the purpose of Layer 2 payload compression and explain how Layer 2 payload compression affects throughput and delay
- Describe the purpose of header compression and explain how header compression affects throughput and delay
- Explain how VoIP packets are susceptible to increased latency when large packets such as FTP transfers traverse slow WAN links
- Explain LFI operation and how LFI reduces the delay and jitter of VoIP packets
- Identify the points in a network where link efficiency mechanisms can most effectively be employed

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Link Efficiency Mechanisms Overview

This topic describes the various LEMs and their functions.

Link Efficiency Mechanisms Overview

- **Link efficiency mechanisms are often deployed on WAN links to increase the throughput and to decrease delay and jitter.**
- **Cisco IOS link efficiency mechanisms include:**
  - **L2 payload compression (Stacker, Predictor, MPPC)**
  - **Header compression (TCP, RTP, class-based TCP, and class-based RTP)**
  - **LFI (MLP, FRF.12, and FRF.11.C)**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-2

While many QoS mechanisms exist for optimizing throughput and reducing delay in network traffic, QoS mechanisms do not create bandwidth. QoS mechanisms optimize the use of existing resources, and enable the differentiation of traffic according to a policy. Link efficiency QoS mechanisms such as payload compression, header compression, and LFI are deployed on WAN links to optimize the use of WAN links.

Payload compression increases the amount of data that can be sent through a transmission resource. Payload compression is primarily performed on Layer 2 frames and therefore compresses the entire Layer 3 packet. The Layer 2 payload compression methods include Stacker, Predictor, and Microsoft Point-to-Point Compression (MPPC).

Compression methods are based on eliminating redundancy. The protocol header is an item of repeated data. The protocol header information in each packet in the same flow does not change much over the lifetime of that flow. Using header compression mechanisms, most header information can be sent only at the beginning of the session, stored in a dictionary, and then referenced in later packets by a short dictionary index. Cisco IOS header compression methods include TCP header compression, Real-Time Transport Protocol (RTP) header compression, class-based TCP header compression, and class-based RTP header compression.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

LFI is a Layer 2 technique in which large frames are broken into smaller, equal-sized fragments, and transmitted over the link in an interleaved fashion with more latency-sensitive traffic flows (like VoIP). Using LFI, smaller frames are prioritized, and a mixture of fragments is sent over the link. LFI reduces the queuing delay of small frames because the frames are sent almost immediately. Link fragmentation, therefore, reduces delay and jitter by expediting the transfer of smaller frames through the hardware transmit (Tx) queue. The LFI methods available include Multilink PPP (MLP), FRF.12, and FRF.11 Annex C.

Layer 2 payload compression squeezes Layer 2 payloads (the entire Layer 3 packet). Layer 2 payload compression increases the throughput and decreases the latency in transmission, because smaller packets (with compressed payloads) take less time to transmit than the larger, uncompressed packets. Layer 2 payload compression is performed on a link-by-link basis.

Header compression methods work by *not* transmitting repeated information in packet headers throughout a session. The two peers on a PPP Layer 2 connection (a dial-up link) agree on session indices which index a dictionary of packet headers. The dictionary is built at the start of every session and is used for all subsequent (non-initial) packets. Only changing, or non-constant, parameters in the headers are actually sent along with the session index.

It is important to note that header compression is performed on a link-by-link basis. Header compression cannot be performed across multiple routers because routers need full Layer 3 header information to be able to route packets to the next hop.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Indexing Operations

Compressed RTP (cRTP) maintains the state for session contexts. A session context is defined by the combination of the IP source and destination addresses, the User Datagram Protocol (UDP) source and destination ports, and the RTP synchronization source (SSRC) field. A compressor implementation might use a hash function on these fields to index a table of stored session contexts. The compressed packet carries a small integer, called the session context identifier, to indicate in which session context that packet should be interpreted. The decompressor can use the context identifier to index its table of stored session contexts directly.

Note Refer to RFC 2507, 2508, and 1144 for more information on header compression operations.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

L2 Payload Compression

This topic describes the purpose of Layer 2 payload compression and how Layer 2 payload compression affects throughput and delay.

The figure shows a basic block diagram of a Layer 2 payload compression method. When a router forwards a packet, the packet is subjected to the Layer 2 compression method after it has been encapsulated at the output. The compression method squeezes the payload of the Layer 2 frame (the entire Layer 3 packet), and transmits the packet on the interface.

Layer 2 payload compression is a CPU-intensive task and can add per-packet compression delay due to the application of the compression method to each frame. The serialization delay, however, is reduced, because the resulting frame is smaller. Serialization delay is the fixed delay that is required to clock the frame onto the network interface. Depending on the complexity of the Layer 2 payload compression algorithm, overall latency might be reduced, especially on low-speed links.

Cisco routers support hardware-assisted compression to reduce the CPU load and the Layer 2 payload compression delay.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: L2 Payload Compression Results

The figure compares three throughput and latency scenarios on a PPP link.

If no compression is used, the throughput is limited by the link bandwidth, and the average delay is influenced by the forwarding or buffering delay, the serialization, and the propagation delay.

If compression is enabled—even if the serialization delay is now shorter because the frame is smaller—the compression or decompression delay may increase the overall latency between the two hops. The perceived throughput is generally increased because the size of the Layer 2 payload is reduced, therefore allowing more Layer 2 frames to be sent through a transmission resource in a given time period. The throughput is limited by the effectiveness of the Layer 2 payload compression algorithm and may be significantly higher than the link bandwidth limit.

If hardware-assisted Layer 2 payload compression is used, the compression or decompression delays may become insignificant compared to forwarding and serialization delays, and overall latency may decrease. The throughput is again limited by the effectiveness of the Layer 2 payload compression method and may be significantly higher than the link bandwidth limit.

Note Layer 2 payload compression configuration will not be covered in this module. Refer to the latest Cisco IOS documentation for configuration details.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Header Compression

This topic describes the purpose of header compression and how header compression affects throughput and delay.

Header compression increases the throughput and reduces the delay by compressing the protocol headers. Header compression is most useful for applications that generate small payloads because the protocol headers of such applications use a significant percentage of bandwidth on a link relative to their payload. Real-time applications typically generate small payloads. Target applications for header compression include Telnet and RTP applications.

TCP and RTP header compression applies to all TCP and RTP flows. For example, if TCP compression is enabled on a link, there is no mechanism to restrict its function to specific application types. TCP header compression for bulk data transfer yields little bandwidth savings. Class-based TCP header compression can be performed only on a certain traffic class, such as the Telnet traffic class.

The header compression algorithm tracks active transport-layer connections over an interface. After the packet has been forwarded, the header compression algorithm compresses the Layer 3 and Layer 4 headers within the frame, and replaces the headers with a session index from the session dictionary (table). Only the non-constant parameters in the headers will be sent along with the session index. The packet is then sent to the output queue, and transmitted to the remote peer. When the remote peer receives the packet, the header is decompressed using the local session table, and passed to the forwarding process.

For example, without RTP header compression, the IP/UDP/RTP header overhead of the voice packet shown in the figure is about 67 percent (40 / 60 x 100 percent). With RTP header compression, the IP/UDP/RTP header can be reduced to 2 or 4 bytes (without and with checksum, respectively) for most packets. Therefore, the IP/UDP/RTP header overhead can be reduced to about 9 percent (2 / 22 x 100 percent) or 17 percent (4 / 24 x 100 percent).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Header Compression Results

The figure compares two throughput and latency scenarios on a PPP link.

If header compression is not used, the throughput is limited by the link bandwidth, and the average delay is influenced only by the forwarding or buffering delay, the serialization, and the propagation delay.

If header compression is enabled, compressing the protocol headers causes the packet to become smaller, therefore allowing more packets to be sent through a transmission resource in a given time period to increase the throughput. Because the packet size is smaller, the serialization delay also becomes smaller, thus reducing the overall delay. Header compression has a low compression delay and a relatively low CPU overhead and is recommended on links slower than 2 Mbps.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Large Packets "Freeze Out" Voice on Slow WAN Links

This topic describes how VoIP packets are susceptible to increased latency when large packets such as FTP transfers traverse slow WAN links.

When considering delay between two hops in a network, queuing delay in a router must be considered because it may be comparable to—or even exceed—the serialization and propagation delay on a link. In an empty network, an interactive or voice session experiences low or no queuing delay, because the session does not compete with other applications on an interface output queue. Also, the small delay does not vary enough to produce considerable jitter on the receiving side.

In a congested network, interactive data and voice applications compete in the router queue with other applications. Queuing mechanisms may prioritize voice traffic in the software queue, but the hardware queue (Tx ring) always uses a FIFO scheduling mechanism. Therefore, after packets of different applications leave the software queue, the packets will mix with other packets in the hardware queue (TxQ), even if their software queue processing was expedited. Therefore, a voice packet may be immediately sent to the hardware TxQ where two large FTP packets may still be waiting for transmission. The voice packet must wait until the FTP packets are transmitted, thus producing an unacceptable delay in the voice path. Because links are used variably, the delay varies with time and may produce unacceptable jitter in jitter-sensitive applications such as voice.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Serialization Delays

	1 Byte	64 Bytes	128 Bytes	256 Bytes	512 Bytes	1024 Bytes	1500 Bytes
56 kbps	143 us	9 ms	18 ms	36 ms	72 ms	144 ms	214 ms
64 kbps	125 us	8 ms	16 ms	32 ms	64 ms	128 ms	187 ms
128 kbps	62.5 us	4 ms	8 ms	16 ms	32 ms	64 ms	93 ms
256 kbps	31 us	2 ms	4 ms	8 ms	16 ms	32 ms	46 ms
512 kbps	15.5 us	1 ms	2 ms	4 ms	8 ms	16 ms	23 ms
768 kbps	10 us	640 us	1.28 ms	2.56 ms	5.1 ms	10.2 ms	15 ms
1536 kbps	5 us	320 us	640 us	1.28 ms	2.56 ms	5.12 ms	7.5 ms

0170_511

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-9

Serialization delay is the fixed delay that is required to clock a voice or data packet onto the network interface. Serialization delay is directly related to the link speed and the size of the packet.

The figure shows the serialization delay as a function of the link speed and packet size.

For example, the serialization delay for a 1500-byte packet over a 56-kbps link will be 214 ms, while the serialization delay is only 7.5 ms over a 1.536-Mbps link for the same 1500-byte packet.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Link Fragmentation and Interleaving

This topic describes LFI operation and how LFI reduces the delay and jitter of VoIP packets.

The use of a hybrid queuing method such as low-latency queuing (LLQ) can provide low latency and low jitter for VoIP packets, while servicing other data packets in a fair manner. But even if VoIP packets are always sent to the front of the software queue, there is still the issue of serialization delay. A large packet may be on its way out of the hardware TxQ, which uses FIFO when a VoIP packet is sent to the front of the software queue. The serialization of the large packet can cause the VoIP packet to wait for a long time before it can be transmitted out.

For example, the serialization delay of a 1500-byte packet over a 56-kbps link will be 214 ms. For VoIP traffic, the maximum recommended one-way, end-to-end delay is 150 ms. Therefore, having a 1500-byte packet ahead of a VoIP packet in the hardware TxQ on a 56-kbps link can cause the end-to-end delay of the voice packet to be over the budget of 150 ms.

The solution to this problem is to fragment the large packets so that they never cause a VoIP packet to wait for more than a predefined amount of time. The VoIP packets must also be allowed to transmit in between the fragments of the larger packets (interleaving), or there will be no point in doing the fragmenting.

When configuring the proper fragment size to use on a link, a typical goal is to have a maximum serialization delay of around 10 to 15 milliseconds. Depending on the LFI mechanisms being configured, the fragment size is either configured in bytes or in milliseconds.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Applying Link Efficiency Mechanisms

This topic describes the points in a network where link efficiency mechanisms can most effectively be employed.

Header compression and LFI are typically configured at the WAN edge for WAN links below T1 or E1 speeds to optimize the use of the WAN link and to prevent long serialization delay.

Layer 2 payload compression is less commonly deployed on WAN links, especially without the use of hardware-assisted payload compression.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- Link efficiency mechanisms (including L2 payload compression, header compression, and LFI) deployed on WAN links can increase throughput and decrease delay and jitter.
- Payload compression uses a compression algorithm to compress the payload of Layer 2 frames.
- Header compression reduces overhead by compressing the IP and upper-layer headers.
- A VoIP packet may be sent to the hardware TxQ, where large FTP packets may still be waiting for transmission. The VoIP packet must wait until the large packets are transmitted, producing an unacceptable delay in the voice path.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-12

Summary (Cont.)

- LFI reduces the delay and jitter of small packets (for example, VoIP) by fragmenting the large packets to allow a VoIP packet to wait no more than a predefined amount of time.
- Header compression and LFI are typically configured at the WAN edge for WAN links below T1 or E1 speeds, to optimize the use of the WAN link and to prevent long serialization delay. Layer 2 payload compression is less commonly deployed on WAN links.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-13

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 2

Configuring Class-Based Header Compression

Overview

Headers exist on almost every communication layer of the OSI stack. When data is sent between workstations, headers will typically be applied at the session, transport, network, and data-link layers. Real-Time Transport Protocol (RTP) is a protocol for the transport of real-time data. RTP includes a data portion and a header portion. The data portion of RTP is a thin protocol that provides support for the real-time properties of applications, such as continuous media, including timing reconstruction, loss detection, and content identification. The header portion of RTP is considerably larger, especially when compared to the payload data that it supports. To avoid the unnecessary consumption of available bandwidth, compressed RTP (cRTP) is used on a link-by-link basis.

TCP header compression is also supported on Cisco routers to reduce the overhead that is associated with TCP and IP headers in TCP/IP packets. TCP header compression is most effective for interactive traffic with small packet size such as Telnet on slow WAN links. The modular quality of service (QoS) command-line interface (CLI), or MQC, method supports class-based RTP and TCP header compression, so that the header compression can be performed on a specific traffic class. This lesson describes class-based RTP and class-based TCP header compression, including configuration and monitoring.

Objectives

Upon completing this lesson, you will be able to configure class-based TCP and class-based RTP header compression to improve bandwidth efficiency and reduce delay. This ability includes being able to meet these objectives:

- Explain the purpose of header compression
- Calculate the overhead that you can save by using class-based TCP header compression
- Calculate the overhead that you can save on VoIP packets using class-based RTP header compression
- Identify the Cisco IOS commands required to configure class-based header compression
- Identify the Cisco IOS commands used to monitor class-based header compression

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Header Compression Overview

This topic describes the purpose of header compression.

Header Compression Overview

- **TCP header compression and class-based TCP header compression:**
  - **Compresses IP and TCP headers**
  - **Used to reduce the overhead of TCP segments**
  - **Most effective on slow links with many TCP sessions with small payloads (for example, Telnet)**
- **RTP header compression and class-based RTP header compression:**
  - **Compresses IP, UDP, and RTP headers**
  - **Used to reduce delay and increase throughput for RTP**
  - **Improves voice quality**
  - **Most effective on slow links**
- **Class-based header compression – Cisco IOS Release 12.2(13)T.**
- **Header compression is enabled on a link-by-link basis.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-2

Three methods are standardized by the Internet Engineering Task Force (IETF) for use with IP protocols:

- **TCP header compression:** Also known as the Van Jacobson or VJ header compression, TCP header compression is used to compress the packet IP and TCP headers over slow links, thus considerably improving the interactive application performance (RFC 1144).
- **RTP header compression:** Used to compress the packet IP, User Datagram Protocol (UDP), and RTP headers, thus lowering the delay for transporting real-time data, such as voice and video, over slower links (RFC 2508).
- **IP header compression:** Not discussed in this lesson. (RFC 2507).

When TCP and RTP header compression is enabled, the compression occurs by default in the fast-switched path or the Cisco Express Forwarding (CEF)-switched path, depending on which switching method is enabled on the interface. Class-based header compression enables RTP or TCP header compression on a per-class basis. This feature was introduced in Cisco IOS Release 12.2(13)T. Decompression is not based on the class map. The receiving end will decompress all packets that come compressed from the other side. Header compression is performed on a link-by-link basis. Header compression cannot be performed across multiple routers because routers need full Layer 3 header information to be able to route packets to the next hop.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: RTP Header Compression

RTP header compression is often implemented on WAN links to reduce the IP, UDP, and RTP header overhead for VoIP packets. Reducing the protocol header overhead results in an increase in throughput and a decrease in serialization delay.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based TCP Header Compression

This topic describes how to calculate the approximate overhead that can be saved by using class-based TCP header compression.

Class-Based TCP Header Compression

- **Most Internet applications use TCP as the transport protocol.**
- **Most of the information in the headers (IP and TCP) is static or predictable throughout the session.**
- **IP (20 bytes) and TCP (20 bytes) use 40 bytes.**
- **TCP header compression can squeeze these two headers into 3 to 5 bytes.**
- **Class-based TCP header compression allows compression on a traffic class.**
- **Class-based TCP header compression is configured via MQC.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-3

In TCP header compression, the IP and TCP headers, which normally use 20 bytes each, are reduced to a session index, and the variable part of the header. With all optimizations, the combined header length of 40 bytes can be reduced to a 3- to 5-byte compressed header.

After you configure a class within a policy map, class-based TCP header compression occurs on a per-class basis. Policy maps are created using the MQC. Attach the policy maps to an interface by using the **service-policy** command. The **service-policy** command gives you the option of specifying either an input service policy (for input interfaces), or an output service policy (for output interfaces). For the TCP header compression feature, only output service policies can be specified.

Enabling TCP header compression on an interface applies header compression to all TCP flows out of the interface. If TCP compression is enabled on a link, there is no mechanism to restrict its function to specific application types. TCP header compression for bulk data transfer (packets with a large payload) yields little bandwidth savings. Using the newer class-based TCP header compression, TCP header compression can be performed only on certain traffic classes; for example, you can only perform TCP header compression on the interactive traffic class where the packet payload sizes are small.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Class-Based TCP Header Compression

Class-Based TCP Header Compression Example

- **Link bandwidth is 64 kbps.**
- **The link is used for a number of interactive TCP sessions.**
- **PPP encapsulation is used.**
- **Average packet size is 5 bytes.**
- **Each segment has 46 bytes of overhead (PPP, IP, and TCP headers).**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-4

The example shows the benefits of TCP header compression on slow links.

A 64-kbps link is used to transport a TCP-based application using PPP as the Layer 2 framing protocol. For the case study application (Telnet), the average packet payload size is 5 bytes. Because PPP has 6 bytes of frame header, the total header overhead is 6 + 20 + 20 = 46 bytes, including the PPP, IP, and TCP headers.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based TCP Header Compression Example (Cont.)

Overhead = 46 / (46 + 5)
Overhead = 90%
Delay = (46 + 5) / 64 kbps * 8
Delay = 6 ms

Overhead = 10 / (10 + 5)
Overhead = 67%
Delay = (10 + 5) / 64kbps * 8
Delay = 2 ms

Payload size	Overhead No comp.	Delay on 64 kbps	Overhead Comp.	Delay on 64 kbps
10	82%	7 ms	50%	2 ms
50	48%	12 ms	17%	7 ms
100	32%	18 ms	9%	13 ms
500	8%	67 ms	2%	62 ms
1500	3%	189 ms	1%	184 ms

0170_084

© 2006 Cisco Systems, Inc. All rights reserved.

The figure shows the packet size before and after TCP header compression. After TCP header compression, the IP and TCP headers are reduced to 4 bytes, resulting in 10 bytes of overall headers. The overhead is reduced from 90 percent to 67 percent when small packets are used. Because of the packet size reduction, the serialization delay decreases from 6 ms to 2 ms on the same 64-kbps link.

The table in the figure shows how TCP header compression impacts performance when different packet sizes are used. TCP header compression is most effective on small packets, and is often used on slow links.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based RTP Header Compression

This topic describes how to calculate the overhead that can be saved on VoIP packets using class-based RTP header compression.

Class-Based RTP Header Compression

- **Voice sessions use RTP.**
- **RTP uses UDP for transport.**
- **Most of the information in the headers (IP, UDP, and RTP) is static throughout the session.**
- **IP (20 bytes), UDP (8 bytes), and RTP (12 bytes) use 40 bytes.**
- **RTP header compression can squeeze these three headers into 2 or 4 bytes.**
- **Class-based RTP header compression allows compression on a traffic class.**
- **Class-based RTP header compression is configured via MQC.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-6

RTP is the standard protocol (RFC 1889) for the transport of real-time data. RTP is intended to provide end-to-end network transport functions for applications that support audio, video, or simulation data over multicast or unicast network services. RTP is used in most VoIP applications to transport voice packets.

RTP includes a data portion and a header portion. The data portion of RTP is a thin protocol that provides support for the real-time properties of applications, such as continuous media, and includes timing reconstruction, loss detection, and content identification. RTP contains a relatively large-sized header. The 12 bytes of the RTP header, combined with 20 bytes of IP header and 8 bytes of the UDP header, create a 40-byte header. For compressed-payload audio applications, the RTP packet typically has a 20- to 160-byte payload, depending on the audio compression codec. Given the size of the IP, UDP, and RTP header combinations, it is inefficient to send the header without compressing it.

To avoid unnecessary consumption of available bandwidth, the RTP header compression feature, cRTP, is used on a link-by-link basis. cRTP can reduce the header from 40 bytes to 2 or 4 bytes (4 bytes with UDP checksum enabled), which significantly reduces delay on slow links.

After you configure a class within a policy map, class-based RTP header compression occurs on a per-class basis. Policy maps are created using the MQC. Attach the policy maps to an interface by using the **service-policy** command. The **service-policy** command gives customers the option of specifying either an input service policy (for input interfaces), or an output service policy (for output interfaces). For the cRTP feature, only output service policies can be specified.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Class-Based RTP Header Compression

Class-Based RTP Header Compression Example

- **Link bandwidth is 64 kbps.**
- **The link is used for VoIP.**
- **PPP encapsulation is used.**
- **G.729 codec is used (8 kbps of voice data, 50 samples per second, 20 bytes per sample).**
- **Each segment has 46 bytes of overhead (PPP, IP, UDP, and RTP headers).**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-7

The example shows the benefits of RTP header compression on slow links.

A 64-kbps link is used to transport VoIP using PPP as the Layer 2 framing protocol.

For the case study application (voice, using the G.729 audio compression codec), the payload size is 20 bytes. Because PPP has 6 bytes of frame header, the total header overhead is 6 + 20 + 8 + 12 = 46 bytes, including the PPP, IP, UDP, and RTP headers.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Class-Based RTP Header Compression Example (Cont.)

Overhead = 46 / (46 + 20) = 70%
Delay = (46 + 20) / 64 kbps * 8 = 8 ms
BW = (46 + 20) * 50 * 8 = 26.4 kbps

2 voice sessions / 64 kbps

Overhead = 10 / (10 + 20) = 33%
Delay = (10 + 20) / 64 kbps * 8 = 4 ms
BW = (10 + 20) * 50 * 8 = 12 kbps

5 voice sessions / 64 kbps

Codec	Voice Bandwidth	RTP Bandwidth	RTP Overhead	CRTP Bandwidth	CRTP Overhead
G.711	64 kbps	82 kbps	22%	68 kbps	3%
G.729	8 kbps	26 kbps	61%	12 kbps	33%

01170_386

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-8

The figure shows the packet size before and after RTP header compression. The IP, UDP, and RTP headers are reduced to 4 bytes, resulting in 10 bytes of overall headers. The overhead is reduced from 70 percent to 33 percent when small packets are used. Because of the packet size reduction, the serialization delay decreases from 8 ms to 4 ms, and the bandwidth that is used to transport a single voice call (using the G.729 codec) is reduced from 26.4 kbps (66 bytes per frame * 50 frames per second * 8 bits per byte) to 12 kbps (30 bytes per frame * 50 frames per second * 8 bits per byte). Therefore, a 64-kbps link can support up to two G.729 voice calls without cRTP, but up to five G.729 voice calls with cRTP.

The table in the figure shows how RTP header compression impacts performance when a different audio codec is used. For the traditional G.711 voice codec, RTP header compression still optimizes its transmission over slow links. However, the difference is more obvious when using advanced, low-bandwidth codecs.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring Class-Based Header Compression

This topic describes the Cisco IOS commands that are required to configure class-based header compression.

Configuring Class-Based Header Compression

```
router(config-pmap-c)#
compression header ip [rtp | tcp ]
```

- **Enables RTP or TCP IP header compression for a specific traffic class.**
- **If the** rtp **or** tcp **options are not specified, both RTP and TCP header compressions are configured.**
- **The number of concurrent compressed connections is automatically determined based on interface bandwidth.**
- **Can be used at any level in the policy map hierarchy configured with MQC.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-9

Configure class-based TCP and RTP header compression within a policy map using the **compression header ip** command.

If you do not specify either RTP or TCP header compression, both RTP and TCP header compressions will be configured.

Note	Header compression is only auto-negotiated on PPP links. On Frame Relay or high-level data link control (HDLC) links, both ends of the links have to match.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Configuring Class-Based TCP Header Compression

Example: Configuring Class-Based TCP Header Compression

```
class-map interactive
 match protocol telnet
!
policy-map cust1
 class interactive
 bandwidth 64
 compression header ip tcp
!
<output omitted>
!
int s0/0
service-policy output cust1
<output omitted>
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-10

In the figure, the compression header ip command is configured to use TCP header compression for a traffic class called “interactive.” The “interactive” traffic class is part of a policy map called “cust1.” This “cust1” policy map is applied to the s0/0 interface in the outbound direction.

This policy provides a minimum bandwidth guarantee of 64 kbps for the “interactive” traffic class and will perform TCP header compression on the “interactive” traffic class (all Telnet packets in this example) leaving the s0/0 interface.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Configuring Class-Based RTP Header Compression

Example: Configuring Class-Based RTP Header Compression

```
class-map voip
 match protocol rtp
!
policy-map cust1
!
class voip
 priority 384
 compression header ip rtp
!
<output omitted>
!
int s0/0
service-policy output cust1
<output omitted>
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-11

In the figure, the compression header ip command has been configured to use RTP header compression for a traffic class called “voip.” The “voip” traffic class is part of a policy map called “cust1.” This “cust1” policy map is applied to the s0/0 interface in the outbound direction.

This policy provides a maximum bandwidth guarantee of 384 kbps for the “voip” traffic class and will perform RTP header compression on the “voip” traffic class (all RTP packets in this example) leaving the s0/0 interface.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring Class-Based Header Compression

This topic describes the Cisco IOS commands that are used to monitor class-based header compression.

Monitoring Class-Based Header Compression

```
router>
show policy-map interface interface-name
```

- **Displays the packet statistics of all classes configured for all service policies on the specified interface**

```
router>show policy-map interface Serial 0/0
Serial0/0
Service-policy output:cust1
Class-map: voip (match-all)
1005 packets, 64320 bytes
30 second offered rate 16000 bps, drop rate 0 bps
Match:protocol rtp
Queueing
        Strict Priority
        Output Queue: Conversation 264
        Bandwidth 384 (kbps) Burst 9600 (Bytes)
        (pkts matched/bytes matched) 1000/17983
        (total drops/bytes drops) 0/0compress:
compress:
header ip rtp
UDP/RTP Compression:
Sent:1000 total, 999 compressed,
41957 bytes saved, 17983 bytes sent
3.33 efficiency improvement factor
99% hit ratio, five minute miss rate 0 misses/sec, 0 max rate 5000 bps
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-12

The **show policy-map interface** command output displays the type of header compression configured (RTP in this example), the interface to which the policy map called “cust1” is attached (serial 0/0), the number of packets sent, the number of packets compressed, the number of bytes saved, and the number of bytes sent.

Other statistical information provided in the output includes the “efficiency improvement factor,” which indicates the percentage of increased bandwidth efficiency as a result of header compression. For example, an efficiency improvement factor of 3.33 means 330 percent efficiency improvement. The “hit ratio” is the percentage of packets that are found in the context database. In most instances, this percentage should be high. The “five-minute miss rate” is the number of traffic flows in the last five minutes that were *not* found in the context database. The rate is the actual traffic rate after the packets are compressed.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **TCP header compression compresses the IP and TCP headers. RTP header compression compresses the IP, UDP, and RTP headers. Header compression reduces delay.**
- **Class-based TCP header compression enables TCP header compression on a per-class basis. TCP header compression can reduce the header from 40 bytes to 3 to 5 bytes.**
- **RTP header compression feature (cRTP) can reduce the header from 40 bytes to a 2- or 4-byte header (4 bytes with UDP checksum enabled), which significantly reduces delay on slow links.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-13

Summary (Cont.)

- **Class-based TCP and RTP header compression is configured within a policy map using the** router(config-pmp-c)# compression header ip **command.**
- **The** show policy-map interface **command output displays the type and statistics of the header compression configured.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-14

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 3

Configuring Link Fragmentation and Interleaving

Overview

Because the hardware transmit queue (TxQ) is a FIFO queue, having a large packet such as an FTP packet in front of a small packet like a VoIP packet can cause excessive delay for the VoIP packet. Link fragmentation and interleaving (LFI) allows the large packets to be fragmented into smaller fragments, and then the small unfragmented packets are interleaved in between the fragments. LFI reduces the delay and jitter of small packets such as VoIP packets over a WAN link.

This lesson describes the configuration and monitoring of two different LFI mechanisms supported in Cisco IOS software, Multilink PPP LFI (MLP LFI) and Frame Relay Fragmentation (FRF.12).

Objectives

Upon completing this lesson, you will be able to configure LFI to improve bandwidth efficiency and reduce delay. This ability includes being able to meet these objectives:

- Identify the different options available for link fragmentation
- Given a list of link speeds and a specific delay requirement, determine the proper fragment size to use at each link speed and the typical delay requirements for VoIP packets
- Identify the Cisco IOS commands required to configure MLP LFI
- Identify the Cisco IOS commands used to monitor MLP LFI
- Explain when FRF.12 can be used and how FRF.12 affects VoIP packets
- Identify the Cisco IOS commands required to configure FRF.12
- Identify the Cisco IOS commands required to monitor FRF.12

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Fragmentation Options

This topic describes the different options that are available for link fragmentation.

Fragmentation Options

Cisco IOS LFI mechanisms include:

- **Multilink PPP with interleaving: PPP links**
- **FRF.12: Frame Relay PVC carrying data traffic, including VoIP over Frame Relay traffic**
- **FRF.11 Annex C: Frame Relay PVC carrying VoFR traffic**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-2

LFI is a Layer 2 technique, in which all Layer 2 frames are broken into small, equal-sized fragments, and transmitted over a link in an interleaved fashion. LFI reduces delay and jitter by expediting transfer of smaller frames through the hardware TxQ.

There are three LFI mechanisms implemented in Cisco IOS software:

- **MLP LFI:** By far the most common and widely used form of LFI
- **FRF.12 Frame Relay LFI:** Used with Frame Relay data connections
- **FRF.11 Annex C LFI:** Used with Voice over Frame Relay (VoFR)

Note For configuration information on FRF.11.C, refer to the latest Cisco IOS documentation.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Serialization Delay and Fragment Sizing

This topic describes the proper fragment size to use on links based on their speed, and identifies the typical delay requirements for VoIP packets.

Serialization Delays

Frame Size

Link Speed	1 Byte	64 Bytes	128 Bytes	256 Bytes	512 Bytes	1024 Bytes	1500 Bytes
56 kbps	143 us	9 ms	18 ms	36 ms	72 ms	144 ms	214 ms
64 kbps	125 us	8 ms	16 ms	32 ms	64 ms	128 ms	187 ms
128 kbps	62.5 us	4 ms	8 ms	16 ms	32 ms	64 ms	93 ms
256 kbps	31 us	2 ms	4 ms	8 ms	16 ms	32 ms	46 ms
512 kbps	15.5 us	1 ms	2 ms	4 ms	8 ms	16 ms	23 ms
768 kbps	10 us	640 us	1.28 ms	2.56 ms	5.1 ms	10.2 ms	15 ms
1536 kbps	5 us	320 us	640 us	1.28 ms	2.56 ms	5.12 ms	7.5 ms

- **For 1500-byte packets, fragmentation is not necessary above T1 (1.5Mbps)**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-3

Serialization delay is the fixed delay required to clock a voice or data packet onto the network interface. Serialization delay is directly related to the link speed and the size of the packet.

The figure shows the serialization delay as a function of the link speed and packet size. For example, the serialization delay for a 1500-byte packet over a 64-kbps link will be:

```
(1500 bytes * 8 bits per byte) / 64 kbps) = 187 ms
```

To ensure low delay and jitter for voice packets on slow links, the recommended standard goal for serialization delay is about 10 to 15 ms. Therefore, on a T1 or higher speed link, LFI is not necessary because the serialization delay for a 1500-byte packet is only 7.5 ms on a T1 link. A T1 link has a bandwidth of 1536 kbps, excluding the T1 framing overhead.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Fragment Size Recommendation for Voice

Link Bandwidth	10 ms	20 ms	30 ms	40 ms	50 ms	100 ms	200 ms
56 kbps	70 bytes	140 bytes	210 bytes	280 bytes	350 bytes	700 bytes	1400 bytes
64 kbps	80 bytes	160 bytes	240 bytes	320 bytes	400 bytes	800 bytes	1600 bytes
128 kbps	160 bytes	320 bytes	480 bytes	640 bytes	800 bytes	1600 bytes	3200 bytes
256 kbps	320 bytes	640 bytes	960 bytes	1280 bytes	1600 bytes	3200 bytes	6400 bytes
512 kbps	640 bytes	1280 bytes	1920 bytes	2560 bytes	3200 bytes	6400 bytes	12800 bytes
768 kbps	1000 bytes	2000 bytes	3000 bytes	4000 bytes	5000 bytes	10000 bytes	20000 bytes
1536 kbps	2000 bytes	4000 bytes	6000 bytes	8000 bytes	10000 bytes	20000 bytes	40000 bytes

Recommendation for Voice (< 15 ms)

017G_513

© 2006 Cisco Systems, Inc. All rights reserved.

To meet the standard serialization delay goal of no more than 10 to 15 ms to ensure low delay and jitter for voice packets, you should configure a fragment size of about 80 bytes per every 64 kbps of the clocking rate for the interface.

Depending on the LFI mechanism, the fragment size is either configured in bytes or in ms. For example, MLP LFI maximum fragment size is configured in ms while FRF.12 maximum fragment size is configured in bytes.

Example: Determining the Proper Fragment Size

The following are example values of fragment size, based on the rule that a fragment size of about 80 bytes per every 64 kbps of the clocking rate should be configured:

- On a 64-kbps link, the proper fragment size to use is 80 bytes.
- On a 128-kbps link, the proper fragment size to use is 160 bytes (80 * 2).
- On a 192-kbps link, the proper fragment size to use is 240 bytes (80 * 3).
- On a 256-kbps link, the proper fragment size to use is 320 bytes (80 * 4).
- And so on.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring MLP with Interleaving

This topic describes the Cisco IOS commands required to configure MLP with interleaving.

Configuring MLP with Interleaving

Configuration steps:

- **Enable MLP on an interface (using a multilink group interface).**
- **Enable MLP interleaving on the multilink interface.**
- **Specify maximum fragment size by setting the maximum delay on the multilink interface.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-5

To configure MLP with interleaving, you must perform these configuration steps:

Step 1 Enable MLP on a PPP interface.

Step 2 On the multilink interface, enable interleaving within MLP.

Step 3 In the multilink interface configuration, specify the maximum fragment size by specifying the maximum desired serialization delay in ms.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring MLP with Interleaving

```
router(config-if)#
ppp multilink
```

- **Enables MLP**

```
router(config-if)#
ppp multilink interleave
```

- **Enables interleaving of frames with fragments**

```
router(config-if)#
ppp multilink fragment delay delay
```

- **Configures maximum fragment delay in ms.**
- **The router calculates the maximum fragment size from the interface bandwidth and the maximum fragment delay.**
- **Fragment size = interface bandwidth * maximum fragment delay**
- **Default maximum fragment delay is 30 ms.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-6

The **ppp multilink** command enables MLP on a PPP interface.

The **ppp multilink interleave** command enables interleaving of fragments within the multilink connection.

The **ppp multilink fragment delay *delay*** command specifies the maximum desired fragment delay for the interleaved multilink connection. The maximum fragment size is calculated from the interface bandwidth and the specified maximum delay. The default is set at 30 ms. To support voice packets, a maximum fragment size of 10 to 15 ms should be used.

If distributed Cisco Express Forwarding (dCEF) is configured on a Versatile Interface Processor (VIP) interface, MLP with interleaving will run in distributed mode on the VIP.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: MLP with Interleaving

MLP with Interleaving Example

```
interface Multilink1
 ip address 172.22.130.1 255.255.255.252
 ppp multilink
 ppp multilink group 1
 ppp multilink fragment delay 10
 ppp multilink interleave
 bandwidth 128
 service-policy output llq-policy
!
interface Serial0/0
 no ip address
 encapsulation ppp
 ppp multilink
 ppp multilink group 1
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-7

The figure shows an example configuration of MLP with interleaving on a multilink group interface.

A non-default maximum desired delay of 10 ms is configured.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring MLP with Interleaving

This topic describes the Cisco IOS commands that are used to monitor MLP with interleaving.

Monitoring MLP Interleaving

```
router>
show interfaces multilink interface-number
```

- **Displays MLP statistics including the number of interleaved frames**

```
router>show interfaces multilink 1
Multilink1 is up, line protocol is up
  Hardware is multilink group interface
  Internet address is 172.22.130.1/30
  MTU 1500 bytes, BW 64 Kbit, DLY 100000 usec,
     reliability 255/255, txload 27/255, rxload 1/255
  Encapsulation PPP, loopback not set
  Keepalive set (10 sec)
  DTR is pulsed for 2 seconds on reset
  LCP Open, multilink Open
  Open: IPCP
  Last input 00:00:03, output never, output hang never
  Last clearing of "show interface" counters 6d00h
  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: weighted fair
  Output queue: 0/1000/64/0/2441 (size/max total/threshold/drops/interleaves)
     Conversations  0/7/16 (active/max active/max total)
     Reserved Conversations 0/0 (allocated/max allocated)
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 7000 bits/sec, 6 packets/sec
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-8

The **show interfaces multilink** command output includes MLP LFI statistics information and indicates whether MLP interleaving is enabled on the interface. Multilink should be in the open state along with Link Control Protocol (LCP) and IP Control Protocol (IPCP).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring MLP Interleaving (Cont.)

```
router#
debug ppp multilink fragments
```

- **Displays information about individual multilink fragments and interleaving events**

```
router#debug ppp multilink fragments
Multilink fragments debugging is on

Mar 17 20:03:08.995: Se0/0 MLP-FS: I seq C0004264 size 70
Mar 17 20:03:09.015: Se0/0 MLP-FS: I seq 80004265 size 160
Mar 17 20:03:09.035: Se0/0 MLP-FS: I seq 4266 size 160
Mar 17 20:03:09.075: Se0/0 MLP-FS: I seq 4267 size 160
Mar 17 20:03:09.079: Se0/0 MLP-FS: I seq 40004268 size 54
Mar 17 20:03:09.091: Se0/0 MLP-FS: I seq C0004269 size 70
Mar 17 20:03:09.099: Se0/0 MLP-FS: I seq C000426A size 70
Mar 17 20:03:09.103: Mu1 MLP: Packet interleaved from queue 24
Mar 17 20:03:09.107: Se0/0 MLP-FS: I seq C000426B size 70
Mar 17 20:03:09.119: Se0/0 MLP-FS: I seq C000426C size 70
Mar 17 20:03:09.123: Mu1 MLP: Packet interleaved from queue 24
Mar 17 20:03:09.131: Mu1 MLP: Packet interleaved from queue 24
Mar 17 20:03:09.135: Se0/0 MLP-FS: I seq C000426D size 70
Mar 17 20:03:09.155: Se0/0 MLP-FS: I seq C000426E size 70
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-9

The **debug ppp multilink fragments** command is a valuable troubleshooting tool when monitoring MLP LFI operations. This command outputs the result of every fragmentation operation, indicating whether the packets are fragmented into correct-sized fragments.

Caution This command should be used with extreme caution in a production environment, because of the amount of output that is created.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

FRF.12 Frame Relay Fragmentation

This topic describes when FRF.12 can be used and how FRF.12 affects VoIP packets.

FRF.12 Frame Relay Fragmentation

FRF.12 specifies fragmentation of Frame Relay data frames:

- **Frame Relay data frames that exceed the specified fragmentation size are fragmented.**
- **Smaller time-sensitive packets can be interleaved.**
- **This is the recommended Frame Relay fragmentation method to be used with VoIP over Frame Relay.**
- **Fragments VoIP over Frame Relay packets if the fragment size is set to a value smaller than the voice packet size.**
- **FRF.12 requires FRTS or DTS.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-10

FRF.12 is the recommended fragmentation to be used with VoIP over Frame Relay while FRF.11 Annex C is the recommended fragmentation to be used with VoFR.

The FRF.12 Implementation Agreement defines FRF.12 fragmentation, and was developed to allow long data frames to be fragmented into smaller pieces and interleaved with real-time frames. In this way, real-time voice and non-real-time data frames can be carried together on lower-speed Frame Relay links without causing excessive delay and jitter to the real-time traffic like VoIP.

Because Frame Relay is a Layer 2 protocol, it has no way to tell which frame contains voice (VoIP) or data. Therefore, Frame Relay will fragment all packets larger than the fragment size into smaller frames, including VoIP packets. In a VoIP over Frame Relay network, it is important to configure the fragment size on the Data-Link Connection Identifier (DLCI) so that VoIP frames will not get fragmented. For example, a G.711 VoIP packet without cRTP is 200 bytes long. For this DLCI, do not set the fragment size to less than 200 bytes.

Frame Relay permanent virtual circuits (PVCs) not configured for VoFR use normal Frame Relay (FRF.3.1) data encapsulation. If fragmentation is turned on for this DLCI, it uses FRF.12 for the fragmentation headers.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

FRF.12 specifies three types of fragmentation, of which only one form is supported in Cisco IOS software:

- **End-to-end:** Supported in Cisco IOS software
  - Packets contain the FRF.12 fragmentation header
  - Occurs at the PVC level
  - LMI packets are not fragmented
- **User-Network Interface (UNI):** Not supported in Cisco IOS software releases
- **Network-to-Network Interface (NNI):** Not supported in Cisco IOS software releases

FRF.12 is configured on a per-PVC basis. If one PVC or subinterface has fragmentation but another does not on the same physical interface, large packets can and will delay smaller packets. For applications where fragmentation is needed, you must turn on fragmentation on all PVCs or subinterfaces that carry large packets. FRF.12 fragmentation is configured within a Frame Relay map class using the **frame-relay fragment** command. The configured map class is then associated with the specific DLCIs. In addition, you must enable Frame Relay traffic shaping (FRTS) on the interface in order for FRF.12 fragmentation to work.

Cisco IOS Release 12.1(5)T introduces a distributed version of FRF.12 Frame Relay fragmentation to use with Distributed Traffic Shaping (DTS).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Configuring FRF.12 Frame Relay Fragmentation

This topic describes the Cisco IOS commands that are required to configure FRF.12.

Configuring FRF.12 Frame Relay Fragmentation

```
router(config)#
map-class frame-relay map-class-name
```

- **Specifies a map class to define QoS values for a virtual circuit**

```
router(config-map-class)#
frame-relay fragment fragment-size
```

- **Enables fragmentation of Frame Relay frames for a Frame Relay map class**
- **Sets the maximum fragment size in bytes**

```
router(config-if)# | (config-subif)# | (config-fr-dlci)#
frame-relay class name
```

- **Associates a map class with an interface, subinterface, or DLCI**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-11

FRF.12 fragmentation is configured within the Frame Relay map class. The **frame-relay fragment** command sets the maximum fragment size in bytes. On an interface, the **frame-relay class** command applies the map class to the interface, subinterface, or a DLCI.

FRF.12 requires FRTS to be enabled.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: FRF.12 Frame Relay Fragmentation

FRF.12 Frame Relay Fragmentation Example

```
interface serial 0/0
  encapsulation frame-relay
  frame-relay traffic-shaping
!
interface serial 0/0.1 point-to-point
 frame-relay interface-dlci 100
  class FRF12
!
map-class frame-relay FRF12
 frame-relay fragment 80
!FRTS parameters
 frame-relay cir 64000
 frame-relay bc 640
 frame-relay mincir 64000
 service-policy output llq-policy
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-13

The figure shows a configuration example where FRF.12 fragmentation is applied to a data Frame Relay circuit configured on the serial 0/0.1 subinterface.

The maximum fragment size is set to 80 bytes. This would be used in a VoIP over Frame Relay environment.

FRTS is enabled on the interface and the FRTS token bucket parameters are configured within the Frame Relay map class. In this figure, FRTS is enabled with a Committed Information Rate (CIR) of 64 kbps, a committed burst (Bc) (normal burst size) of 640 bits, and uses weighted fair queuing (WFQ) as the shaping queue.

Note The purpose of fragmentation is to decrease serialization delays so that large data frames cannot block small voice packets for excessive periods of time. Both voice and data packets are subject to the shaping rate. FRTS token bucket operation is based on burst of credits in bits. If a large data packet gets through the shaper and uses up all the credits, the voice packets will be delayed until more credit becomes available. But if only a small fragment of a data packet can get through the shaper, the voice packet will be delayed for less time. So with FRTS and FRF.12, the bottleneck is at the shaper, not at the port. Therefore, the fragment size should be determined based on the CIR and not the physical interface clocking rate.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

FRTS doesn't take into account the flags and cyclic redundancy checks (CRCs) in the shaping algorithm. Because a flag and CRC need to be added for every fragment, on lower-speed links (768 kbps and below), you should shape to a value slightly lower than the actual CIR to account for the flags and CRCs. For example, on a 128-kbps Frame Relay PVC, the recommended frame-relay CIR value to use will be 124,872 bps with a Bc of 1250 bits per committed time window (Tc).

PVC Speed	Fragment Size (for 10ms delay)	Recommended CIR	Recommended Bc (CIR/100)
56 kbps	70 bytes	52968 bps	530 bits
64 kbps	80 bytes	60952 bps	610 bits
128 kbps	160 bytes	124872 bps	1250 bits
256 kbps	320 bytes	252832 bps	2530 bits
512 kbps	640 bytes	508816 bps	5090 bits
768 kbps	960 bytes	764940 bps	7560 bits

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring FRF.12 Frame Relay Fragmentation

This topic describes the Cisco IOS commands that are required to monitor FRF.12.

Monitoring FRF.12 Frame Relay Fragmentation

```
router>
show frame-relay fragment [interface interface [DLCI]]
```

- **Displays information about the Frame Relay fragmentation**

```
router>show frame-relay fragment
interface        dlci  frag-type    frag-size  in-frag    out-frag   dropped-frag
Serial0/0.1      100   end-to-end   80         0          0          0
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-14

The **show frame-relay fragment** command displays information about the FRF.12 Frame Relay fragmentation process. The fragment type will always display end-to-end because this is the only type presently supported on Cisco IOS software. In addition to fragment type, the fragment size in bytes and associated DLCI is displayed.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Monitoring FRF.12 Frame Relay Fragmentation (Cont.)

```
router>
show frame-relay pvc [interface interface] [dlci]
```

- **Displays statistics about PVCs for Frame Relay interfaces**

```
router>show frame-relay pvc 100

PVC Statistics for interface Serial0/0 (Frame Relay DTE)
DLCI = 100, DLCI USAGE = LOCAL, PVC STATUS = INACTIVE, INTERFACE = Serial0/0.1
<output omitted>
  Current fair queue configuration:
   Discard     Dynamic         Reserved
   threshold  queue count  queue count
    64            16                  0
  Output queue size 0/max total 600/drops 0
  fragment type end-to-end fragment size 80
  cir 64000     bc   2600      be 0         limit 325    interval 40
  mincir 32000     byte increment 320   BECN response no  IF_CONG no
  frags 0         bytes 0         frags delayed 0         bytes delayed 0

  shaping inactive
  traffic shaping drops 0
```

© 2006 Cisco Systems, Inc. All rights reserved.

The **show frame-relay pvc** command output includes settings related to the FRF.12 fragmentation process. This output shows the fragment size (80 bytes in this example) used on the Frame Relay PVC. The fragment type is end-to-end because Cisco IOS software currently only supports end-to-end FRF.12 fragmentation.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **Cisco LFI options are Multilink PPP with interleaving, FRF.12, and FRF.11 Annex C.**
- **Serialization delay is a function of the link speed and packet size. To ensure low delay and jitter for voice packets on slow links, the recommended standard goal for serialization delay is to not exceed 10 to 15 ms.**
- **The** ppp multilink **command enables MLP on a PPP interface. The** ppp multilink interleave **command enables interleaving of fragments within the multilink connection. The** ppp multilink fragment-delay **command specifies the maximum desired fragment delay for the interleaved multilink connection.**
- **The** show interface multilink **command output includes MLP LFI statistics information and indicates whether MLP interleaving is enabled on the interface.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-16

Summary (Cont.)

- **FRF.12 specifies fragmentation of Frame Relay data frames when they exceed a specified size or when there are smaller time-sensitive packets that can be interleaved, and will fragment VoIPs if fragment size is set to a smaller value than the VoIP size.**
- **The** frame-relay fragment **map-class command sets the maximum fragment size in bytes. On an interface, the** frame-relay class **command applies the map class to the interface, subinterface, or a DLCI.**
- **The** show frame-relay fragment **command displays information about the FRF.12 Frame Relay fragmentation statistics.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-17

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Summary

This topic summarizes the key points discussed in this module.

Module Summary

- **Link efficiency mechanisms are used to increase the perceived throughput and decrease the latency on WAN links.**
- **Class-based TCP and class-based RTP header compression enable header compression on specific traffic classes using the MQC method.**
- **LFI improves the serialization delay for packets leaving the router and prevents the voice packets from waiting on large data packets to be processed.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—8-1

The different link efficiency mechanisms available in Cisco IOS software include class-based TCP and RTP header compression, Layer 2 payload compression, MLP link fragmentation and interleaving (MLP LFI), and Frame Relay Fragmentation (FRF.12). You can use these QoS mechanisms on slower WAN links, such as 768 kbps and slower, to improve the link efficiency by increasing the throughput and decreasing the overall delay.

Class-based TCP and RTP header compressions are configured using the MQC method to reduce the protocol header overhead. Header compression is most effective for packets with small payload size.

Layer 2 payload compression is not used much because of the software compression delay overhead. Using a hardware compression module can help reduce the software compression delay.

For MLP LFI is supported on a leased-line interfaces, while FRF.12 (for VoIP) and FRF.11 Annex C (for VoFR) are supported on Frame Relay interfaces. LFI is used to reduce the serialization delay of large packets. Serialization delay is based on the link speed and packet size. To ensure high voice quality, the recommended serialization delay is 10 to 15 ms maximum. As a general rule, a fragment size of 80 bytes for every 64 kbps of link speed is recommended.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

References

For additional information, refer to these resources:

- For more information on link efficiency mechanisms, refer to *Link Efficiency Mechanisms Overview* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt6/qcflem.htm
- For more information on Layer 2 payload compression, refer to *Understanding Data Compression* at this URL: http://www.cisco.com/warp/public/116/compress_overview.html
- For more information on TCP and RTP header compression, refer to these RFCs:
  - RFC 2508 (RTP header compression)
  - RFC 1144 (TCP header compression)
  - RFC 2507 (IP header compression)
- For more information on FRF.12 and DTS configuration, refer to *Frame Relay Traffic Shaping With Distributed QoS on the Cisco 7500 Series* at this URL: http://www.cisco.com/warp/public/105/distribfrts.html#topic2%20URL
- For information on FRF.11 Annex C, refer to *VoFR Encapsulation and Fragmentation* at these URLs:
  - http://www.cisco.com/warp/public/788/voip/vofr_encap_frag_5733.html
  - http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121newft/121limit/121e/121e2/vofvip/vofr3vip.htm
- For more information on Cisco class-based TCP and RTP header compression, refer to *Class-Based RTP and TCP Header Compression* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t13/fthdrcmp.htm#wp1044256

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check

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

Q1) With Layer 2 payload compression, what can be done to improve the compression/decompression delay of the router? (Source: Understanding Link Efficiency Mechanisms)

A) enable CEF switching
B) enable fast switching
C) use the Stacker or Predictor compression algorithm
D) use hardware-assisted compression

Q2) Which two of these statements are true, based on the following **show** output? (Choose two.) (Source: Configuring Class-Based Header Compression)

```
wg6r1#show policy-map interface s0/1
Serial0/1
Service-policy output: cust1
Class-map: voip (match-all)
300 packets, 202035 bytes
5 minute offered rate 2000 bps, drop rate 0 bps
Match: protocol rtp
Queuing
Strict Priority
Output Queue: Conversation 264
Bandwidth 384 (kbps) Burst 9600 (Bytes)
(pkts matched/bytes matched) 232/139975
(total drops/bytes drops) 15/15795
compress:
header ip rtp
UDP/RTP compression:
Sent: 285 total, 284 compressed,
9086 bytes saved, 176014 bytes sent
1.5 efficiency improvement factor
99% hit ratio, five minute miss rate 0 misses/sec, 0 max
rate 2000 bps
```

A) Both class-based TCP and RTP header compression are enabled for the VoIP traffic class.
B) IP payload compression is enabled for the VoIP traffic class.
C) LLQ is enabled for the VoIP traffic class.
D) Class-based RTP header compression is enabled for all RTP traffic.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q3) Which two of these statements are correct regarding TCP and RTP header compression? (Choose two.) (Source: Configuring Class-Based Header Compression)

A) Hardware-assisted header compression is required to reduce the header compression delay.
B) TCP header compression compresses both the IP and TCP headers.
C) RTP header compression compresses the IP, UDP, and RTP headers.
D) TCP and RTP header compression can compress the respective protocol headers down to 10 bytes.

Q4) TCP header compression is most effective on which of these applications? (Source: Configuring Class-Based Header Compression)

A) VoIP
B) FTP
C) video streaming
D) Telnet
E) TFTP

Q5) RTP header compression is most effective on which of the following applications? (Source: Configuring Class-Based Header Compression)

A) FTP
B) VoIP
C) TFTP
D) Telnet
E) HTTP

Q6) Which three statements about FRF.12 and FRF.11 Annex C are correct? (Choose three.) (Source: Configuring Link Fragmentation and Interleaving)

A) FRF.12 fragmentation is used on DLCIs configured for VoIP.
B) FRF.12 fragmentation is used on DLCIs configured for VoFR.
C) FRF.11 Annex C fragmentation is used on DLCIs configured for VoIP.
D) FRF.11 Annex C fragmentation is used on DLCIs configured for VoFR.
E) FRF.12 fragmentation requires FRTS to be enabled.
F) Cisco IOS software supports UNI and NNI FRF.12 fragmentation, and not end-to-end fragmentation.

Q7) When configuring Multilink PPP with interleaving, where is the fragment size configured? (Source: Configuring Link Fragmentation and Interleaving)

A) The fragment size in milliseconds is configured under the physical serial interfaces with PPP encapsulation and Multilink PPP enabled.
B) The fragment size in milliseconds is configured under the logical multilink interface.
C) The fragment size in bytes is configured under the map class.
D) The fragment size in bytes is configured within the policy map.

Q8) Which two factors will influence the serialization delay? (Choose two.) (Source: Configuring Link Fragmentation and Interleaving)

A) link speed
B) speed of light in the media
C) router CPU processing power
D) packet size

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q9) To ensure good voice quality, what is the recommended fragment size? (Source: Configuring Link Fragmentation and Interleaving)

A) 80 bytes per every 64 kbps of the clocking, which will result in a 10-ms serialization delay
B) 40 bytes per every 64 kbps of the clocking, which will result in a 20-ms serialization delay
C) 20 bytes per every 64 kbps of the clocking, which will result in a 10-ms serialization delay
D) 120 bytes per every 64 kbps of the clocking, which will result in a 150-ms serialization delay

Q10) From the **show interface multilink** output, which state(s) should be in the open state to indicate proper Multilink PPP operation over an IP interface? (Source: Configuring Link Fragmentation and Interleaving)

A) LCP, Multilink, and IPCP
B) LCP, IPCP, LFICP
C) LCP and MLP
D) LCP and IPCP

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check Answer Key

Q1)	D
Q2)	C, D
Q3)	B, C
Q4)	D
Q5)	B
Q6)	A, D, E
Q7)	B
Q8)	A, D
Q9)	A
Q10)	A

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module 9

QoS Best Practices

Overview

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

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

This module describes the implementation of the DiffServ model in service provider and enterprise networks. The first lesson describes baseline traffic classifications. The second lesson describes a DiffServ implementation in a typical enterprise campus and service provider network.

Module Objectives

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

- Describe the set of classification practices that most closely represent Cisco QoS best practices
- Describe the set of QoS mechanisms used to implement Cisco end-to-end QoS best practices in a typical enterprise network connected thru a service provider that is providing Layer 3 IP services

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 1

Understanding Traffic Classification Best Practices

Overview

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

Objectives

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

- Define some of the key QoS best-practices recommendations
- Explain the QoS requirements of the different application types
- Describe typical enterprise traffic classes and the delay, jitter, packet loss, and bandwidth requirements of each traffic class
- Map different enterprise application types into the appropriate PHB
- Describe how to map different enterprise traffic classes into appropriate service provider traffic classes

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS Best Practices

This topic describes some of the QoS best-practices recommendations.

QoS Best Practices Enterprise Networks

- **Networks must provide predictable, measurable, and sometimes guaranteed services.**
- **Achieve the required QoS by managing the delay, delay variation (jitter), bandwidth, and packet-loss parameters on a network.**
- **Classify and mark traffic as close to the source as possible.**
- **Use NBAR to discover the applications on the network.**
- **Ensure that real-time traffic gets priority with minimal delay.**
- **Ensure that business-critical traffic is correctly serviced.**
- **Ensure that scavenger traffic (for example, file sharing) does not consume too much valuable bandwidth.**
- **Use link efficiency techniques on WAN links.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-2

An IP network forms the backbone of any successful organization. IP networks transport applications and data, including high quality video and delay-sensitive data (such as real-time voice). These applications stretch the IP network capabilities and resources, but also complement, add value to, and enhance every business process. IP networks must provide secure, predictable, measurable, and sometimes guaranteed services. Achieving the required QoS by managing the delay, delay variation (jitter), bandwidth, and packet-loss parameters on an IP network is the key to a successful end-to-end business networking solutions.

QoS is the set of techniques and tools that are used to manage network resources. The figure lists some of the key QoS best-practices recommendations.

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

- IP source and destination addresses
- Layer 4 protocol and port numbers
- Incoming interface
- MAC address
- IP Precedence

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

- DSCP value
- Layer 2 information (Frame-Relay discard eligible [DE] bits, Ethernet 802.1P bits)
- The Cisco value-added mechanism Network-Based Application Recognition (NBAR)

It is best practice to classify and mark the traffic as close to the source of the traffic as possible.

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS Best Practices
Service Provider Networks

- **Use DiffServ backbone**
- **Typically define 3 to 4 service classes at the network core:**
  - **Controlled latency (EF) class**
  - **One or two controlled load classes**
  - **Best-effort class**
- **Typically over provision the controlled latency class by a factor of 2**
- **May have more service classes at the network edge:**
  - **Example: management class**
- **Use LLQ or MDRR (GSR), WRED**
- **When policing and re-marking excess traffic, re-mark to the same AF class with higher drop probability:**
  - **Example: Re-mark AF 31 exceeding traffic to AF 32 (do not re-mark AF 31 to AF 21)**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-3

In an IP service provider environment, the recommended architecture includes the deployment of a DiffServ backbone with a different overprovisioning ratio for different traffic classes. Most service providers typically classify their customer traffic into three or four traffic classes (such as controlled latency, controlled load, and best effort). At the network edge, the service provider may define additional traffic classes, such as, for example, a management traffic class for all network management traffic (such as Simple Network Management Protocol [SNMP] and Telnet traffic).

In the service provider core, where all the links are very high-speed, the typical QoS mechanisms required are LLQ/CBWFQ or MDRR (on a Gigabit Switch Router - GSR) and weighted random early detection (WRED).

At the ingress to the service provider core, the customer traffic is typically policed by the service provider to ensure that the rate of customer traffic does not exceed the contractual rate. When re-marking the exceeding customer traffic, re-mark the DSCP value to the same Assured Forwarding (AF) class, but increase the drop probability. For example, re-mark AF 31 exceeding traffic to AF 32 or AF 33; do not re-mark AF 31 exceeding traffic to AF 21 or AF 11.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

A set of QoS tools is needed to manage the delay, delay variation (jitter), bandwidth, and packet-loss parameters on a network. These QoS tools include QoS classification, scheduling, and provisioning tools.

Only after traffic is identified can traffic policies be applied. The first element of a QoS policy is to identify the traffic to be treated differently. Classification must be consistent throughout the enterprise for it to be effective. QoS classification tools mark a packet with a specific priority. This marking is determined by examining the following:

- Layer 2 parameters such as 802.1Q class of service (CoS) bits, MAC address, and Multiprotocol Label Switching (MPLS) experimental bits
- Layer 3 parameters such as IP precedence, DSCP, and source or destination IP address
- Layer 4 parameters such as TCP or User Datagram Protocol (UDP) ports

At the network edge, QoS markings may be accepted or rejected. This is referred to as the trust boundary. If packets are entering the network from a source that is trusted to set the packet marking properly, the packets are coming from a trusted device and can be left marked as-is. If the packets are entering the network from a device that cannot be trusted to properly mark packets, these packets must be re-marked. This is the location of the trust boundary.

The QoS scheduling tools are used to determine how a packet exits a node. Whenever packets enter a device faster than they can exit the device (as with a speed mismatch), then a point of congestion can occur. Devices have buffers that allow for scheduling of higher-priority packets to exit sooner than lower-priority packets, which is commonly called queuing. Queuing is used only when a device is experiencing congestion, and is bypassed when the congestion clears.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The QoS provisioning tools include traffic-policing and traffic-shaping tools and link efficiency mechanisms. Traffic policers and shapers are the oldest form of QoS mechanisms. These tools have the same objectives: they must identify and respond to traffic violations. Policers and shapers usually identify traffic violations in a similar manner; however, the main difference between policers and shapers is the manner in which they respond to violations, as follows:

- A policer usually drops or remarks excess traffic.
- A shaper delays excess traffic using buffer to hold packets and to shape the flow when the data rate is higher than expected.

Link efficiency mechanisms are typically used on slow-speed WAN links to improve their throughput and to reduce their delays. Link efficiency mechanisms include LFI, payload compression, and header compression.

Implementing QoS is a means to optimize bandwidth, but not a blanket substitute for bandwidth itself. When the network is faced with ever-increasing congestion, a certain point is reached where QoS alone will not solve bandwidth requirements. At such a point, nothing short of additional bandwidth will suffice.

Example: Cisco IOS QoS Tools Summary

In this course, you have learned about the various QoS tools available within Cisco IOS software.

The following list groups some of these tools into the proper categories:

- Classification tools:
  - Class-based marking
  - NBAR
  - QoS Policy Propagation on Border Gateway Protocol (BGP), or QPPB
  - QoS preclassify
- Scheduling tools:
  - CBWFQ
  - LLQ
  - WRED
- Provisioning Tools:
  - Class-based shaping
  - Class-based policing
  - MLP LFI
  - FRF.12
  - cRTP

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Voice, Video, and Data QoS Requirements

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

Voice quality is directly affected by these QoS quality factors:

- **Loss:** Loss causes voice clipping and skips. The industry standard coded algorithms that are used in Cisco digital signal processors (DSPs) can correct for up to 30 ms of lost voice. For example, if a 20 ms sample of voice payload is used per Voice over IP (VoIP) packet, only a single voice packet can be lost during any given time. If two successive voice packets are lost, the 30-ms correctable window is exceeded and voice quality begins to degrade.
- **Delay:** Delay can cause voice quality degradation if the delay is above 200 ms. The International Telecommunication Union (ITU) standard for VoIP (G.114) states that a 150-ms one-way delay budget is acceptable for high voice quality.
- **Jitter:** There are dynamic and adaptive jitter buffers within Cisco IP telephony devices. However, these can usually only compensate for 20 to 50 ms of jitter, or delay variations.

To determine the bandwidth requirement per each VoIP call, several factors must be considered. These factors include the codec used, the sampling rate, and the Layer 2 protocol. Each VoIP call also requires 150 bps of bandwidth for the voice control (call-signaling) traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Provisioning for Voice: VoIP Bandwidth Reference Tables

CODEC	Packetization Interval	Voice Payload in Bytes	Packets per Second	Bandwidth per Conversion
G.711	20 ms	160	50	80 kbps
G.711	30 ms	240	33	74 kbps
G.729A	20 ms	20	50	24 kbps
G.729A	30 ms	30	33	19 kbps

A more accurate method for provisioning is to include the Layer 2 overhead into the bandwidth calculations:

CODEC	801.Q Ethernet + 32 L2 Bytes	MLP + 13 L2 Bytes	Frame Relay + 8 L2 Bytes	ATM + Variable L2 Bytes (Cell Padding)
G.711 at 50 pps	93 kbps	86 kbps	84 kbps	106 kbps
G.711 at 33 pps	83 kbps	78 kbps	77 kbps	84 kbps
G.729A at 50 pps	37 kbps	30 kbps	28 kbps	43 kbps
G.729A at 33 pps	27 kbps	22 kbps	21 kbps	28 kbps

017G_S16

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-6

When addressing the bandwidth requirement for voice traffic, keep the following in mind:

- The required bandwidth per call depends on the codec, the packetization interval, and the Layer 2 protocol.
- The bandwidth required for the voice control (signaling) traffic is about 150 bps per call.

The bandwidth consumed by VoIP streams is calculated by adding the packet payload (in bits) and all the headers (in bits), then multiplying the total number of bits per packet by the per-second packet rate. This does not take into account the effect of any link efficiency tools such as RTP header compression and LFI. The graphic illustrates the VoIP bearer traffic bandwidth requirements for different codecs with different sampling rates. For a more accurate method for VoIP bandwidth provisioning, also include the Layer 2 overhead (that is, preambles, headers, flags, cyclic redundancy checks (CRCs), ATM cell-padding, and others) in bandwidth calculations.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: G.711 Voice Bearer Bandwidth Requirement Calculation

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

```
G.711             = 160 bytes payload size

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

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

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

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


The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS Traffic Requirements: Videoconferencing

- Latency < 150 ms*
- Jitter < 30 ms*
- Loss < 1%*
- Minimum priority bandwidth guarantee required is:
  - Video stream + 20%
  - For example, a 384-kbps stream would require 460 kbps of bandwidth

*One-way requirements

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-7

Videoconferencing has the same latency, jitter, and packet-loss 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. Because of its bursty nature, videoconferencing has these two unique best-practice requirements in provisioning for strict-priority bandwidth:

- The LLQ should be provisioned to the stream rate plus 20 percent.
- The LLQ burst parameter should be provisioned to 30 KB per 384-kbps stream.

Compared to videoconferencing, streaming video applications have more lenient QoS requirements. Video streaming applications are more delay insensitive and are largely jitter insensitive due to application buffering. When provisioning for streaming video applications, take into account the video content distribution requirements. Video file distribution is similar to FTP traffic in nature and can have a major impact on network performance because of the large file size. Content distribution traffic should be managed to avoid impacting the network.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Calculating the Bandwidth Requirement for a 384-kbps Videoconference Stream

This example shows how to calculate the bandwidth requirement for a 384-kbps videoconferencing stream:

```
384 kbps + (20% x 384 kbps)   = 384 kbps + 76.8 kbps
                              = 460.8 kbps
```

Because bandwidth requirements vary greatly from application to application, and even between versions of the same applications, it is not possible to provide a blanket rule for provisioning data bandwidth.

Traffic analysis and lab testing are required to ascertain bandwidth requirements for data applications.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Provisioning for Data: General Principles

- **Profile applications to their basic network requirements.**
- **Do not overengineer provisioning. Use no more than four to five traffic classes for data traffic:**
  - **Mission-Critical: Locally defined critical applications**
  - **Transactional: ERP, SAP, Oracle**
  - **Best-Effort: E-mail, unspecified**
  - **Less-than-Best-Effort (Scavenger): Peer-to-peer applications**
- **Do not assign more than three applications to mission-critical or transactional classes.**
- **Only group applications with common characteristics together into the same class.**
- **Most applications fall under best effort; make sure that adequate bandwidth is provisioned for this default class.**
- **Seek executive endorsement of relative ranking of application priority prior to rolling out QoS policies for data.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-9

For data classification, a relative-priority model of no more than four to five traffic classes has been proven to work well in most enterprise environments, such as, for example, a mission-critical, transactional, best-effort (default), and less-than-best-effort (scavenger) class. Each traffic class can be assigned a minimum bandwidth guarantee using CBWFQ. An example of using a relative-priority model between the different data traffic classes is as follows:

- The mission-critical traffic class gets a minimum of 50 percent of the remaining bandwidth after the LLQ is serviced.
- The transactional traffic class gets a minimum of 20 percent of the remaining bandwidth after the LLQ is serviced.
- The best-effort traffic class gets a minimum of 25 percent of the remaining bandwidth after the LLQ is serviced.
- The less-than-best-effort traffic class gets a minimum of 5 percent of the remaining bandwidth after the LLQ is serviced.

It is counterproductive to use too many discrete classes for data traffic, because the more classes that are defined, the less the distinction between service-levels. Additionally, if too many applications are assigned to the mission-critical class, the overall effectiveness of QoS provisioning is lessened. Taken to an extreme, if all applications are assigned to the mission-critical class, the end result is the same as when no QoS is provisioned at all. Mission-critical applications should be limited to those that directly contribute to the core operations of the business. These are usually highly interactive and are therefore sensitive to delay and loss. Examples of mission-critical applications include ERP applications, such as Oracle, SAP, and PeopleSoft, and proprietary applications that were designed in-house.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

When classifying the different network applications into traffic classes, try to group only applications with common characteristics and QoS requirements together into the same traffic class. For example, because interactive applications and bulk transfer applications have very different bandwidth and delay requirements, interactive and bulk traffic should not be classified into the same traffic class.

Example: RCL Enterprise

In this example, the RCL enterprise network administrator used NBAR and determined that the major sources of network traffic are from the following applications:

- Structured Query Language (SQL)
- VoIP
- FTP
- HTTP
- Domain Name System (DNS)
- Exchange
- Napster

The RCL senior management team has determined that only SQL and HTTP traffic are mission-critical to the company.

RCL has decided to implement four traffic classes as follows:

- A real-time traffic class for VoIP traffic—all RTP traffic
- A mission-critical data traffic class for mission-critical traffic—SQL and HTTP traffic
- A bulk traffic class for business traffic that is not mission-critical—FTP, DNS, Exchange
- A best-effort traffic class for all best-effort traffic—Napster and all other traffic

The RCL network administrator has decided not to implement a scavenger traffic class to limit the Napster traffic at this time, but is constantly monitoring the Napster traffic. In the future, if more bandwidth is required for the business-related applications, a scavenger traffic class will be added and the Napster traffic will be classified into this scavenger traffic class.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS Requirements Summary

This topic describes typical enterprise traffic classes and the delay, jitter, packet-loss, and bandwidth requirements of each enterprise traffic class.

QoS Requirements Summary

	Voice	Videoconference	Bulk Data (FTP)	Mission-Critical Data
Bandwidth	Low to Moderate	Moderate	Moderate to High	Low to Moderate
Drop-Sensitive	High	High	Low	Moderate to High
Delay-Sensitive	High	High	Low	Low to Moderate
Jitter-Sensitive	High	High	Low	Low to Moderate

017G_515

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-10

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

Example: QoS Requirements of the Major Applications Category

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

Mission-critical applications, such as ERP, have different bandwidth requirements, depending on the application and the particular version of the application. In general, mission-critical applications are more drop-sensitive, delay-sensitive, and jitter-sensitive than bulk traffic.

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Traffic Classification

This topic maps different enterprise application types into the appropriate per-hop behavior (PHB).

Traffic Classification and Scheduling

1. **Identify the network traffic and its QoS requirements.**
2. **Classify the network traffic into the appropriate traffic classes and mark the network traffic as close to the source as possible.**
3. **Define the scheduling policy for each traffic class.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-11

After the QoS requirements for the different applications running on the network have been determined, the next step is to determine how to classify the traffic into multiple service classes. It is best practice to classify and mark traffic as close to the source of the traffic as possible. At the enterprise campus LAN, classification and marking is typically done at the access or distribution layer. Both Cisco IOS routers and IOS Catalyst switches have class-based marking capabilities. NBAR is also supported on Cisco IOS routers and some IOS Catalyst switches (within software).

After the traffic classification requirements have been determined, the next step is to establish the proper scheduling policy for each traffic class. For example, at the WAN edge router where the traffic exits the enterprise campus LAN to enter the WAN, it is best practice to implement LLQ and CBWFQ to provide congestion management.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Enterprise WAN Edge Bandwidth Allocation Example

The amount of maximum bandwidth to be allocated for the LLQ class depends on many factors such as the router platform, the traffic profiles, and QoS features enabled.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-12

The following design principles apply when determining bandwidth provisioning requirements for combining voice, video, and data over a single WAN link:

- Video requires 460 kbps of LLQ for each 384 kbps stream (stream-rate plus 20 percent).
- VoIP also requires LLQ bandwidth. The actual amount of bandwidth depends on the number of IP Phones, codecs used, sampling rates, and Layer 2 protocol.
- The ratios of voice and video within the LLQ will vary, depending on the voice and video traffic requirements.
- As a general rule, the total LLQ provisioning is recommended to be ≤ 33 percent of the link capacity. This maximum number varies depending on many factors, such as the router platform and the traffic profiles.
- In a relative-priority model, mission-critical traffic should have greater bandwidth guarantees than the other data classes.
- Because most traffic may end up in the default (best-effort) class, ensure that there is enough bandwidth left over for the best-effort traffic class.
- By default, the sum of all bandwidth guarantees should be ≤ 75 percent of the link bandwidth. For example, if the link bandwidth is 10 Mbps, the maximum available bandwidth will be 7.5 Mbps.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: LLQ Bandwidth Allocation

The example in the figure shows an LLQ bandwidth allocation between six different traffic classes:

- **VoIP bearer traffic class:** The VoIP bearer traffic class has a maximum bandwidth guarantee of 17 percent of the link bandwidth.
- **Video traffic class:** The video traffic class has a maximum bandwidth guarantee of 16 percent of the link bandwidth.
- **VoIP control traffic class:** The VoIP control traffic class has a minimum bandwidth guarantee of 2 percent of the link bandwidth. Newer versions of Cisco CallManager may require additional bandwidth. Be sure to reference the latest Cisco CallManager IP Telephony Design Guides (at www.cisco.com/go/srnd) for the latest recommendations.
- **Mission-critical data class:** The mission-critical traffic class has a minimum bandwidth guarantee of 25 percent of the link bandwidth.
- **Transactional data class:** The transactional data traffic class has a minimum bandwidth guarantee of 15 percent of the link bandwidth.
- **Best-effort data (default) class:** The best-effort data (default) traffic class has a minimum bandwidth guarantee of 25 percent of the link bandwidth.

Note In this example, the maximum reservable bandwidth on the link is set to 0 percent, so that up to 100 percent of the link bandwidth can be guaranteed between the different classes.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS Baseline Classification Summary

Application	L3 Classification IPP	PHB	DSCP	L2 CoS
Routing	6	CS6	48	6
Voice	5	EF	46	5
Videoconferencing	5	AF41	34	4
Streaming Video	4	CS4	32	4
Mission-Critical Data	3	AF31	26	3
Call-Signaling	3	CS3	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

0170_451

© 2006 Cisco Systems, Inc. All rights reserved.

The figure illustrates a QoS baseline recommendation for traffic classification and markings. In this example, the enterprise traffic is classified into 11 traffic classes including the following:

- Five data classes (mission-critical, transactional, bulk, scavenger, best effort)
- One class for network management traffic
- One class for routing protocol traffic
- One class for voice traffic
- One separate class for voice call-signaling traffic
- One class for videoconferencing traffic
- One class for streaming video traffic

In the figure, the recommended IP precedence value, the PHB, the DSCP value, and the Layer 2 CoS markings are shown for each traffic class. The markings of some of the traffic classes are as follows:

- Voice is marked with DSCP 46 and CoS 5.
- Voice call-signaling traffic is marked with Class Selector (CS3) and CoS 3.
- Videoconferencing (AF41) is marked differently than streaming video (CS4) because the QoS requirements are different for videoconferencing than for streaming video.
- Mission-critical data is marked with AF31 and CoS 3.
- A scavenger class is used and is marked with CS1 and CoS 1.
- Routing traffic is marked with CS6 and CoS 6. Cisco IOS software assigns an IP Precedence of 6 to routing protocol packets on the control plane.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS Baseline Classification Summary (Cont.)

Application	L3 Classification			L2
	IPP	PHB	DSCP	CoS
Routing	6	CS6	48	6
Voice	5	EF	46	5
Videoconferencing	5	AF41	34	4
Streaming Video	4	CS4	32	4
Mission-Critical Data	3	-	25	3
Call-Signaling	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

0170_452

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-14

A marking of CS3 is recommended for voice call-signaling traffic. However, some older versions of Cisco IP Phones and other Cisco IP telephony devices originally marked voice call-signaling traffic to AF31. A reevaluation of the relevant standard (RFC 2597, an Assured Forwarding PHB) led Cisco to change its call-signaling marking to CS3. A Class-Selector PHB is better suited to the needs of call signaling, given that an Assured Forwarding PHB allows for markdown and increased drop-preference of in-transit call-signaling packets, which is undesired considering that these flows form the control-plane of the IP Telephony infrastructure.

For enterprises migrating to CS3 marking for call-signaling, both AF31 and CS3 should be reserved for the voice call- signaling traffic class. Furthermore, mission-critical data should be marked as DSCP 25 instead of AF31, as shown in the previous figure, until migration is complete.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: LLQ Example on the Enterprise WAN Edge Router

WAN Edge LLQ Configuration

```
class-map match-all ROUTING
 match ip dscp cs6
class-map match-all VOICE
 match ip dscp ef
class-map match-all VIDEO-CONF
 match ip dscp af41
class-map match-all STREAM-VIDEO
 match ip dscp cs4
class-map match-all MISSION-CRIT
 match ip dscp 25
class-map match-any VOICE-CONTROL
 match ip dscp cs3
 match ip dscp af31
class-map match-all TRANSACT
 match ip dscp af21
class-map match-all NETWORK-MGMT
 match ip dscp cs2
class-map match-all BULK
 match ip dscp af11
class-map match-all SCAVENGER
 match ip dscp cs1
```

```
policy-map WAN-EDGE
 class ROUTING
  bandwidth percent 3
 class VOICE
  priority percent 18
 class VIDEO-CONF
  priority percent 15
 class STREAM-VIDEO
  bandwidth percent 10
class MISSION-CRIT
  bandwidth percent 12
  random-detect dscp-based
class VOICE-CONTROL
  bandwidth percent 2
class TRANSACT
  bandwidth percent 8
  random-detect dscp-based
class NETWORK-MGMT
  bandwidth percent 2
class BULK
  bandwidth percent 4
  random-detect dscp-based
 class SCAVENGER
  bandwidth percent 1
 class class-default
  bandwidth percent 25
  random-detect dscp-based
```

If the default class is allocated a bandwidth, WFQ cannot be enabled for the traffic within the default class. This is true for all platforms except the Cisco 7500 (and soon the Cisco 7200).

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-15

The figure illustrates an example of an enterprise WAN-edge router configuration using LLQ with class-based WRED on certain data traffic classes.

This assumes the markings for the different traffic classes are already done at the access or at the distribution layer within the campus network.

Currently (except for the Cisco 7500 router platform) all traffic classes except for the default traffic class support only FIFO queuing within the class. Check http://www.cisco.com for the latest information on weighted fair queuing (WFQ) support within each traffic class.

On all platforms, the default traffic class can support either FIFO or WFQ within the class. But if the default traffic class is allocated a minimum bandwidth as shown in the figure, WFQ will not be supported in the default traffic class. The only current exception is for the Cisco 7500 series platforms. In this case, the default traffic class will support only FIFO queuing.

Note In this example, the maximum reservable bandwidth on the link is set to 0 percent, so that up to 100 percent of the link bandwidth can be guaranteed between the different classes.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Enterprise to Service Provider QoS Class Mapping

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

The figure illustrates how the different enterprise traffic classes can be mapped into the four traffic classes offered by a service provider. In the example, there is no streaming video traffic. The following four traffic classes are offered by the service provider:

- **Controlled Latency:** A traffic class for all real-time traffic. The controlled latency class admits packets marked with CS5 and Expedited Forwarding (EF).
- **Controlled Load 1:** A traffic class for all highly bursty traffic. The controlled load 1 class admits packets marked with CS3, AF31, and CS6.
- **Controlled Load 2:** A traffic class for the less bursty traffic. The controlled load 2 class admits packets marked with AF21.
- **Best Effort:** A traffic class for the best-effort traffic. All other traffic is assigned to the best-effort class.

The figure also shows an example of bandwidth allocation between the four traffic classes where 35 percent of the link bandwidth is allocated for the controlled latency traffic class, 43 percent for controlled load 1 traffic class, 15 percent for controlled load 2 traffic class, and 7 percent for best-effort traffic class.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Example: Remapping Enterprise-Managed CE Traffic Classes into Traffic Classes Offered by Service Provider

The figure illustrates a sample customer edge (CE) to provider edge (PE) router configuration using LLQ with class-based WRED on certain data traffic classes.

Class-based markings are also used to re-mark the enterprise markings into the markings that are expected by the service provider, as follows:

- Videoconferencing traffic is re-marked from AF41 to CS5 and has a maximum bandwidth guarantee of 13 percent of the link bandwidth.
- Mission-critical traffic is re-marked from DSCP 25 to AF31 and has a minimum bandwidth guarantee of 25 percent of the link bandwidth.
- Voice control (signaling) traffic is re-marked from AF31 and CS3 to CS5 and has a maximum bandwidth guarantee of 2 percent of the link bandwidth.
- Transactional data traffic is re-marked from AF21 to CS3 and has a minimum bandwidth guarantee of 13 percent of the link bandwidth.
- Network management traffic is re-marked from CS2 to CS3 and has a minimum bandwidth guarantee of 2 percent of the link bandwidth.
- Bulk data traffic is re-marked from AF11 to AF21 and has a minimum bandwidth guarantee of 15 percent of the link bandwidth.
- Scavenger data traffic is re-marked from CS1 to DSCP 0 and has a minimum bandwidth guarantee of 1 percent of the link bandwidth.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

- Routing traffic is not re-marked and has a minimum bandwidth guarantee of 3 percent of the link bandwidth.
- Voice bearer traffic is not re-marked and has a maximum bandwidth guarantee of 20 percent of the link bandwidth.
- The default class is not re-marked and has a minimum bandwidth guarantee of 6 percent of the link bandwidth.

Note In this example, the maximum reservable bandwidth on the link is set to 0 percent, so that up to 100 percent of the link bandwidth can be guaranteed between the different classes.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **QoS Best Practices is achieved by:**
  - **Managing delay, delay variation (jitter), bandwidth, and packet loss**
  - **Classifying and marking traffic as close to the source as possible**
  - **Using NBAR**
  - **Ensuring that real-time traffic gets priority with minimal delay**
  - **Ensuring that business-critical traffic is correctly serviced**
  - **Ensuring that scavenger traffic consumes only a small amount of bandwidth**
  - **Using link efficiency techniques**
- **QoS has specific requirements for enterprise traffic classes: voice, videoconferencing, bulk data, and mission-critical data.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-18

Summary (Cont.)

- **Each enterprise traffic class has delay, jitter, packet loss, and bandwidth requirements.**
- **Specific design principles apply when determining bandwidth provisioning requirements for combining voice, video, and data over a single WAN link.**
- **Different enterprise traffic classes must be mapped into the specific traffic classes offered by the service provider.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-19

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 2

Deploying End-to-End QoS

Overview

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

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

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

Objective

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS Service Level Agreements

This topic describes QoS SLAs and provides some SLA examples.

QoS Service Level Agreements

- **QoS SLAs provide contractual assurance for meeting the different traffic QoS requirements.**
- **QoS SLAs typically provide contractual assurance for parameters such as:**
  - **Delay (fixed and variable)**
  - **Jitter**
  - **Packet loss**
  - **Throughput**
  - **Availability**
- **QoS SLAs are a key differentiator for service providers.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-2

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Enterprise Network with Traditional Layer 2 Service

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

© 2006 Cisco Systems, Inc. All rights reserved.

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Enterprise Network with IP Service

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

© 2006 Cisco Systems, Inc. All rights reserved.

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Service Provider SLA Example

	Controlled Latency	Controlled Load	Best Effort
Delay (40ms)	90%	75%	50%
Jitter (2 ms)	90%	75%	50%
Packet Loss (0.5%)	90%	75%	50%

0177G_510

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

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-7

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

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Deploying End-to-End QoS

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Enterprise Campus QoS Implementations

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

IP telephony, videoconferencing, e-learning, and mission-critical data applications are becoming more common in enterprise networks. QoS functions such as classification, scheduling, and provisioning are required within the campus to manage the switch output buffers to minimize packet loss, delay, and jitter. Some of the general guidelines when implementing campus QoS include the following:

- Use a hierarchical design model for the campus network design, consisting of three layers: access layer, distribution layer, and core layer.
- Use multiple queues on the transmit interfaces to minimize the potential for dropped or delayed traffic caused by transmit buffer congestion.
- Give real-time voice and video traffic priority over data traffic.
- Establish proper trust boundaries. For example, at the access layer switches, only trust the IP Phone CoS marking, not the PC CoS marking.
- Classify and mark the traffic as soon as possible. For example, use class-based markings at the access layer switches to classify and mark the traffic coming into the network. NBAR can be used on the access or distribution switches to classify traffic if the switch supports NBAR.
- At the access or distribution layers, use class-based policing to limit any undesired non-business-related excess traffic.
- Do not classify or mark packets at the core layer, because this slows down traffic.
- Do most of the complex QoS configurations, such as traffic shaping, header compression, link fragmentation and interleaving (LFI), LLQ, WRED, and re-marking of traffic, at the WAN CE router.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The figure illustrates the typical QoS configurations required at the access, distribution, and core layer switches. Typically, the high-speed core layer does not require any QoS mechanisms because all the links are very high-speed.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Catalyst 3550 Access Switch

- 1P3Q2T (recommended) or 4Q2T
- WRR default
  - Each queue is given a relative weight with one priority queue
- Default values are as follows:

CoS Value	CoS Priority Queues
0, 1	1
2, 3	2
4, 5	3
6, 7	4 (priority queue)

- Move CoS 5 traffic to queue 4
- Trust Cisco IP Phone on access links
- Trust DSCP or CoS on uplinks
- CoS-to-DSCP mappings
- Classify and mark mission-critical traffic using ACLs

© 2006 Cisco Systems, Inc. All rights reserved.

For the Catalyst 3550 switches, the recommended configuration for the transmit interface is one priority queue and three queues, each with a single drop threshold.

Scheduling is performed using WRR, in which each queue is given a relative weight while the priority queue is serviced exhaustively. The default CoS to queue mapping is as follows:

- Frames marked with a CoS 6 and 7 will go into queue 4 (priority queue for the Catalyst 3550).
- Frames marked with a CoS 4 and 5 will go into queue 3.
- Frames marked with a CoS 2 and 3 will go into queue 2.
- Frames marked with a CoS 0 and 1 will go into queue 1.

When servicing voice traffic, CoS 5 frames should be configured to go to the priority (expedite) queue. By default, Cisco IP Phones forward voice traffic with an 802.1Q priority of 5 (CoS 5).

At the access layer switch, configure the trust boundary to trust only the CoS marking on the Cisco IP Phone, not the CoS marking on the PC. Also configure the trust boundary to trust the DSCP or CoS marking from the distribution layer switches, depending on the capabilities of the distribution layer switch that is attached.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

For Layer 2-to-Layer 3 QoS mappings, the Catalyst 3550 has default CoS-to-DSCP mappings. However, you can manually configure these default mappings to override the default mappings. The default CoS-to-DSCP map is as follows:

CoS-to-DSCP Default Mapping

Marker	Value							
CoS Values	0	1	2	3	4	5	6	7
DSCP Values	0	8	16	24	32	40	48	56

As a general rule, it is best to classify and mark the traffic as close to the source as possible. Therefore, at the access layer switches, you can configure class-based markings to classify and mark the traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Catalyst 3550 Access Switch Configuration

```
mls qos
!
mls qos map cos-dscp 0 8 16 26 34 46 48 56
!
class-map match-all voice-bearer
 match access-group 101
class-map match-all voice-control
 match access-group 102
class-map match-all mission-critical
 match access-group 103
!
policy-map mark
 class voice-bearer
  set ip dscp ef
 class voice-control
  set ip dscp 26
 class mission-critical
  set ip dscp af31
!
access-list 101 permit udp any any range 16384 32767
access-list 102 permit tcp any any range 2000 2002
access-list 102 permit tcp any any eq 1720
access-list 102 permit tcp any any range 11000 11999
access-list 102 permit udp any any eq 2427
access-list 103 permit <Mission-Critical Traffic>
```

```
!
interface GigabitEthernet0/12
 description Uplink to Distribution
 no ip address
 mls qos trust dscp
 priority-queue out
 wrr-queue cos-map 4 5
!
interface FastEthernet0/1
 description to IP Phone
 cdp enable
 no ip address
 service-policy input mark
 mls qos trust cos
 mls qos trust device cisco-phone
 switchport priority extend cos 0
 switchport voice vlan 111
 switchport access vlan 11
 priority-queue out
 wrr-queue cos-map 4 5

<output omitted>
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-13

The figure shows a sample QoS configuration on a Catalyst 3550 access switch.

When classification of traffic by the access layer switch is required, the Catalyst 3550 provides a powerful set of features for classifying traffic as the traffic enters the network. This configuration example uses three class maps to classify traffic into three classes (voice-bearer, voice-control, and mission-critical) using three extended IP access control lists (ACLs). The access ports are configured to trust only the CoS markings from the Cisco IP Phone; all other traffic will be classified and marked using the traffic policy called “mark.” This traffic policy is used to mark the voice-bearer class with a PHB of expedited forwarding (EF), the voice-control (call-signaling) class with a DSCP of 26, and the mission-critical class with a DSCP of AF 31. This traffic policy is then applied to the switch access ports.

Note If you are using the Catalyst 2950, the ACLs are limited to match a single TCP or User Datagram Protocol (UDP) port. For example, the Catalyst 2950 ACLs cannot be used to match against a UDP or TCP port range for matching VoIP RTP bearer and control traffic.

The uplink switch ports are configured to trust the DSCP markings from the distribution layer switches. Layer 3-capable distribution layer switches are being used in this topic.

The Catalyst 3550 supports multiple VLANs on the access port to which the Cisco IP Phone is attached. The **switchport voice vlan** interface configuration command is used to configure voice VLAN on the port. In this example, all the voice traffic from the access port will belong to VLAN 111 while the data traffic will belong to VLAN 11.

To ensure EF for the voice-bearer traffic, all CoS 5 traffic is configured to go into queue 4 (the priority queue [PQ]) along with the CoS 6 and CoS 7 traffic.

The default CoS-to-DSCP mapping is also changed so that CoS 3 (voice control) maps to AF31 (26), CoS 4 (videoconferencing) maps to AF41 (34), and CoS 5 (voice bearer) maps to DSCP EF (46).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS in Catalyst 4000: Distribution (SUPIII)

- **4 queues (1P3Q2T)**
  - **Configurable PQ for queue 3**
  - **CoS value 4, 5 selects queue 3**
- **Trust DSCP or CoS**
- **Default mapping from CoS to DSCP and DSCP to CoS**
- **Class-based policing to rate-limit traffic**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-14

The Catalyst 4500 has a total of four queues. You can configure the link bandwidth to be shared differently among the four transmit queues by assigning the minimum guaranteed bandwidth for each transmit queue. By default, all four queues are scheduled in a round-robin manner, with each queue having 25 percent of the link bandwidth. You can configure the transmit queue 3 on each port with strict priority. When transmit queue 3 is configured with higher priority, packets in transmit queue 3 are scheduled ahead of packets in other queues. The default DCSP-to-transmit queue map (which you can manually configure to override the default) is as follows:

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

At the distribution switch, configure the trust boundary to trust the DSCP or CoS marking from the core and access layer switches. For Layer 2-to-Layer 3 and Layer 3-to-Layer 2 QoS mappings, the Catalyst 4500 has default CoS-to-DSCP and DSCP-to-CoS mappings. However, you can manually configure these default mappings to override the default mappings. The Catalyst 4500 has these CoS-to-DSCP and DSCP-to-CoS mappings:

CoS-to-DSCP Default Mapping

Marker	Value							
CoS Values	0	1	2	3	4	5	6	7
DSCP Values	0	8	16	24	32	40	48	56

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

DSCP-to-CoS Default Mapping

Marker	Value							
DSCP Values	0 - 7	8 - 15	16 - 23	24 - 31	32 - 39	40 - 47	48 - 55	56 - 63
CoS Values	0	1	2	3	4	5	6	7

You can also implement class-based policing to rate-limit certain traffic classes.

Catalyst 4000 Distribution Switch Configuration

```
class-map citrix-server
 match access-group 1
!
policy-map police-citrix
 class citrix-server
  police 48000 8000 exceed-action drop
!
interface GigabitEthernet4/1
 description to access switch
 service-policy in police-citrix
 qos trust dscp
 tx-queue 3
  priority high
!
interface GigabitEthernet4/2
 description to core
 qos trust dscp
 tx-queue 3
  priority high
!
access-list 1 permit host <citrix-server IP address>
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-15

The figure shows a sample QoS configuration on a Catalyst 4500 distribution switch.

This configuration uses a class map to classify the Citrix server traffic into the citrix-server traffic class.

A traffic policy called “police-citrix” is used to police the Citrix traffic to a committed information rate (CIR) of 48,000 bps with a committed burst (Bc) of 8000 bits. All excess traffic is dropped.

The downlink and uplink ports are configured to trust the DSCP markings from the access and core layer switches.

The transmit queue 3 is configured as the high-priority queue.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WAN Edge (CE/PE) QoS Implementations

This topic describes some of the best-practice QoS implementations and configurations on WAN CE and PE routers.

On a WAN link between CE and PE, LLQ or CBWFQ, traffic shaping, compressed Real-Time Transport Protocol (cRTP), and LFI are typically required.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

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

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Traffic Leaving Service Provider Network

- Service provider enforces SLA using the output QoS policy on PE.
- Output policy uses queuing, dropping, and, optionally, shaping.
- May require LFI or cRTP.
- No input QoS policy on CE needed.

- Service provider enforces SLA using the output QoS policy on PE.
- Output policy uses queuing, dropping, and, optionally, shaping.
- May require LFI or cRTP.
- Input QoS policy on CE irrelevant.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-19

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Traffic Leaving Service Provider Network (Cont.)

CE Input Policy	PE Output Policy
<Not needed>	LLQ
	WRED
	[Shaping]
	[LFI or cRTP]

CE Input Policy	PE Output Policy
<Irrelevant>	LLQ
	WRED
	[Shaping]
	[LFI or cRTP]

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-20

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Managed CE with Three Service Classes Example

- **The service provider in this case study offers managed customer edge service with three service classes:**
  - **Premium (VoIP): Maximum bandwidth guarantee, low latency, no loss**
  - **Business: Minimum bandwidth guarantee, low loss**
  - **Class-Default: No guarantees (best effort)**
- **Most DiffServ deployments use a proportional differentiation model:**
  - **Rather than allocate absolute bandwidths to each class, service provider adjusts relative bandwidth ratios between classes to achieve SLA differentiation:**
    - **VoIP serviced with priority up to a max. percentage of bandwidth, for example, 25%**
    - **Business class gets majority of remaining bandwidth, for example, 75%**
    - **Best effort gets whatever is left, for example, 25% of remaining bandwidth**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-21

In the example, the service provider implements an IP DiffServ backbone and offers these three traffic classes with different SLAs for each:

- **Premium (VoIP):** The premium traffic class is intended for voice traffic. This class has a maximum bandwidth limit, a low latency, and no loss guarantee.
- **Business:** The business traffic class is intended for mission-critical traffic. This class has a minimum bandwidth and a low loss guarantee.
- **Class-Default:** The default (best effort) traffic class is intended for all other traffic. This class has no guarantee.

When implementing LLQ or CBWFQ, the bandwidth guarantees can be specified in kbps, in the percent of the available bandwidth, or in the percent of the remaining available bandwidth. Most DiffServ deployments use a proportional differentiation model, in which the bandwidth guarantees are configured as a percentage instead of a fixed kbps.

For example, with three traffic classes, the bandwidth allocation can be divided as follows:

- The premium (VoIP) class LLQ can have a maximum bandwidth guarantee of 25 percent of the link bandwidth.
- The business (mission-critical) class can have a minimum bandwidth guarantee of 75 percent of the remaining bandwidth after the LLQ is serviced.
- The default (best-effort) class can have a minimum bandwidth guarantee of whatever is left, in this case, 25 percent of the remaining bandwidth after the LLQ is serviced.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

WAN Edge Design

- **VoIP**
  - Packet marked EF class and sent to LLQ
  - Max bandwidth = 25% of CIR
  - Policed
  - Excess dropped
  - VoIP signaling shares the LLQ with VoIP traffic
- **Business data**
  - Packets marked AF 31 class and sent to CBWFQ
  - Allocated 75% of remaining bandwidth once LLQ has been serviced
  - Policed
  - Exceeding or violating traffic re-marked
  - WRED configured to optimize TCP throughput
- **Best Effort**
  - Best-effort class sent to CBWFQ
  - Whatever is left = 25% of remaining bandwidth after LLQ has been serviced
  - WRED configured to optimize TCP throughput

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-22

The figure shows the required PHB for each of the three traffic classes supported by the service provider.

For the premium (VoIP) traffic class, the VoIP packets will be marked with EF and will go into the LLQ with these parameters:

- The LLQ will be policed and have a maximum bandwidth of 25 percent of the CIR.
- All excess traffic will be dropped.
- The call-signaling traffic will share the LLQ with the VoIP bearer traffic.

For the business (mission-critical) traffic class, the business class packets will be marked with AF 31 and will go into the CBWFQ with these class parameters:

- The class will be policed and have a minimum bandwidth guarantee of 75 percent of the remaining available bandwidth.
- All exceeding and violating traffic will be re-marked, then sent.
- WRED will be used on this traffic class to optimize TCP throughput.

For the default (best-effort) traffic class, the best-effort class packets will be marked with DSCP of 0, with these class parameters:

- The class is not policed and has a minimum bandwidth guarantee of 25 percent of the remaining available bandwidth.
- WRED will be used on this traffic class to optimize TCP throughput.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The figure shows the QoS configurations on the managed CE router outbound interface to implement the required QoS policy required for each of the three service provider traffic classes. The traffic policy called “OUT-POLICY” is configured to provide the LLQ or CBWFQ, and WRED. Each traffic class bandwidth guarantee is configured using a percentage rather than a fixed bandwidth in kbps.

Both the enterprise voice bearer and voice control traffic will be matched using access-list 101 and will be serviced by the LLQ with a maximum bandwidth guarantee of 25 percent of the link bandwidth.

The enterprise mission-critical traffic will be matched using access-list 102 and has a minimum guaranteed bandwidth of 75 percent of the remaining available bandwidth after the LLQ has been serviced.

All other traffic from the enterprise will be classified into the default class and has a minimum guaranteed bandwidth of 25 percent of the remaining available bandwidth after the LLQ has been serviced.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

In the example, the CE and PE link is a Frame Relay link and traffic shaping is implemented on the PVC using Frame Relay traffic shaping (FRTS).

FRTS is configured using a Frame Relay map class with a CIR of 256 kbps, a Bc of 2560 bits, a Be of 0 (no bursting), and a minCIR of 256 kbps. The CIR is the rate at which you want to normally send when there is no congestion. The CIR needs to be the remote end-link speed or the actual CIR of the VC. The Bc is the amount you will send per time interval. The CIR and Bc will be used to compute a committed time window (Tc), where Tc = Bc / CIR. For FRTS, CLI will only allow Bc values that would result in a Tc that is between 10 ms and 125 ms. A recommended value for Tc is 10 ms. To get a Tc of 10 ms, the Bc should be set to 1/100 of the CIR. The minCIR parameter is required, because Cisco IOS will only allow 75 percent of the minCIR to be reserved by the QoS policy applied to the FRTS map class. If the minCIR is not configured, it will default to 50 percent of the CIR, which is not desired.

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

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

Note	If the Frame Relay PVC is carrying VoIP traffic, the recommendation is to set the minCIR = CIR and not use adaptive traffic shaping, because by the time the backward explicit congestion notification (BECN) is received by the sender, drops are already happening on the Frame Relay switch.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CE-to-PE QoS for Frame Relay Access PE Inbound

```
class-map match-all PREMIUM
 match ip dscp ef
class-map match-all BUSINESS
 match ip dscp af31
!
policy-map IN-POLICY
  class PREMIUM
   police percent 25
    conform-action transmit
    exceed-action drop
  class BUSINESS
   police percent 38
    conform-action transmit
    exceed-action set-dscp-transmit af32
    violate-action set-dscp-transmit af33
  class class-default
   set ip dscp 0
!
interface Serial0/0.1 point-to-point
 ip address 10.32.14.2 255.255.255.0
 frame-relay interface-dlci 16
 class FR-class
!
map-class frame-relay FR-class
 frame-relay adaptive-shaping becn
 frame-relay cir 256000
 frame-relay bc 2560
 frame-relay be 0
 frame-relay mincir 128000
 service-policy output IN-POLICY
 frame-relay fragment 320
 frame-relay ip rtp header-compression
```

© 2006 Cisco Systems, Inc. All rights reserved.

The figure shows the QoS configurations on the ingress PE router inbound interface to implement the required QoS policy that is required for each of the three service provider traffic classes.

In this case, a traffic policy called “IN-POLICY” is configured to provide the required class-based policing. For the premium class, the rate limit is set to 25 percent of the link bandwidth. All exceeding premium-class traffic is dropped. For the business class, the rate limit is set to 38 percent of the link bandwidth. All exceeding and violating business-class traffic is re-marked with a higher drop probability, then sent. The default class is not policed.

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

CE-to-PE QoS for Frame Relay Access PE Inbound (Cont.)

```
class-map match-all PREMIUM
  match ip dscp ef
class-map match-all BUSINESS
  match ip dscp af31
!
policy-map IN-POLICY
 class PREMIUM
  police percent 25
   conform-action set-mpls-exp-transmit 5
   exceed-action drop
 class BUSINESS
  police percent 38
   conform-action set-mpls-exp-transmit 4
   exceed-action  set-mpls-exp-transmit 3
   violate-action set-mpls-exp-transmit 2
 class class-default
   set mpls experimental 0
!
interface Serial0/0.1 point-to-point
 ip address 10.32.14.2 255.255.255.0
 frame-relay interface-dlci 16
 class FR-class
!
map-class frame-relay FR-class
 frame-relay adaptive-shaping becn
 frame-relay cir 256000
 frame-relay bc 2560
 frame-relay be 0
 frame-relay mincir 128000
 service-policy output IN-POLICY
 frame-relay fragment 320
 frame-relay ip rtp header-compression
```

© 2006 Cisco Systems, Inc. All rights reserved.

The figure is similar to the previous figure except that the service provider core is now Multiprotocol Label Switching (MPLS) instead of Frame Relay. The differences are the markings within the traffic policy. In this case, a traffic policy called “IN-POLICY” is configured to provide the required class-based policing. The premium-class and business-class traffic is policed and then marked with these MPLS experimental bits:

- All conforming premium traffic is marked with MPLS Experimental 5, then sent.
- All exceeding premium traffic is dropped.
- All conforming business traffic is marked with MPLS Experimental 4, then sent.
- All exceeding business traffic is marked with MPLS Experimental 3, then sent.
- All violating business traffic is marked with MPLS Experimental 2, then sent.
- The best-effort class traffic is not policed but just marked with MPLS Experimental 0.

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Service Provider Backbone QoS Implementations

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

Service Provider Backbone QoS Implementation

- **Marking, policing, and shaping should be done at the edges of the network.**
- **Queuing and dropping are done in the core based on packet marking done at the edge.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-27

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

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Service Provider Backbone

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

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-28

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

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

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

What Are the Benefits of Using a DiffServ Backbone?

- **DiffServ allows support of multiple classes of traffic with different underprovisioning and overprovisioning ratios per class of service.**
- **Maximum potential economic benefit of DiffServ is when traffic requiring the highest SLA represents a minor proportion of the overall capacity.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-29

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

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

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

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

PE-to-P QoS PE Outbound

```
class-map match-all PREMIUM
 match ip dscp ef
!
class-map match-all BUSINESS
  match ip dscp af31 af32 af33
!
policy-map OUT-POLICY
  class PREMIUM
    priority percent 25
  class BUSINESS
    bandwidth percent remaining 75
    random-detect dscp-based
  class class-default
    bandwidth percent remaining 25
    random-detect dscp-based
!
interface POS1/0
 ip address 10.150.1.1 255.255.255.0
 service-policy output OUT-POLICY
```

PE Outbound

LLQ

WRED

© 2006 Cisco Systems, Inc. All rights reserved.

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

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

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

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

No inbound policy is required on the P router.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

P-to-P QoS P Outbound

```
class-map match-all PREMIUM
 match ip dscp ef
!
class-map match-all BUSINESS
  match ip dscp af31 af32 af33
!
policy-map OUT-POLICY
  class PREMIUM
    priority percent 25
  class BUSINESS
    bandwidth percent remaining 75
    random-detect dscp-based
  class class-default
    bandwidth percent remaining 25
    random-detect dscp-based
!
interface POS1/0
 ip address 10.160.1.1 255.255.255.0
 service-policy output OUT-POLICY
```

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-31

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

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

No inbound policy is required on the receiving P router.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

No inbound policy is required on the egress PE router.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

No inbound policy is required on the CE router.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- An SLA stipulates the delivery and pricing of numerous service levels. SLAs cover an assortment of data services such as Frame Relay, leased lines, Internet access, web hosting, etc.
- There are QoS requirements for different traffic types. Both the enterprise and the service provider must implement the proper QoS mechanisms to provide end-to-end QoS.
- General guidelines for enterprise QoS implementations:
  - Robust switching design
  - Buffer management
  - Multiple queues
  - Voice traffic always gets highest priority
  - Trust Cisco IP Phone CoS setting
  - Classify and mark traffic as close to source as possible
  - Class-based policing in effect

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-35

Summary

- On a CE-to-PE WAN link, LLQ or CBWFQ, traffic shaping, cRTP, and LFI are typically required.
- The service provider IP core is used to provide high-speed packet transport. Therefore, all markings, policing, and shaping should be performed only at the PE router on the PE-to-CE link and not at the core. Using the DiffServ model, only the edge requires complex QoS policy. At the core, only queuing and WRED are required. The operation of queuing and WRED will be based on the markings done at the edge (PE).

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-36

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Lesson 3

Providing QoS for Security

Overview

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

Objective

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

- Describe the business security threat
- Describe the threat of DoS and worm attacks
- Describe the impact a worm can have on the infrastructure of campus and remote sites
- Describe tools and tactics used to increase security
- Describe Control Plane Policing
- Describe Scavenger Class
- Describe the characteristics of using NBAR for known-worm mitigation
- Describe the impact on Internet worms when QoS security policies are in place

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Business Security Threat

This topic describes business security threats.

In recent years, there has been an increase in the frequency of DoS and worm attacks, and an increase in their sophistication.

DoS or worm attacks fall into two main classes: spoofing attacks and flooding attacks. In spoofing attacks, the attacker pretends to provide a legitimate service but provides false information to the requestor. A flooding attack, on the other hand, is an attack which generates and propagates so much traffic that service resources are overwhelmed. Spoofing attacks are best mitigated by authentication and encryption technologies, while flooding attacks can be mitigated using QoS technologies.

Most flooding attacks target PCs and servers, which in turn, when infected, then target other PCs and servers, multiplying traffic flows which result in these resources becoming overwhelmed.

One way that QoS could mitigate a flooding issue would be to use QoS policies to react to out-of-profile network behavior by way of access layer policers. Policers would not impact normal traffic flows or in-profile network behavior, but rather would be configured to start dropping packets aggressively when abnormal traffic flows start to occur. The policers would not know that an attack is in progress—the policers would only know that traffic is at a level at which packets need to be dropped in order to keep the QoS within the network.

One best practice is to use a scavenger-class strategy to mitigate attacks. For example, by using access layer policers, traffic can be marked down to this class and re-marked so that packet drops occur more aggressively. This policy would impact PC and server traffic that may peak to out-of-profile, but that is exactly the purpose of the markdown.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Increasing Occurrence of Network Attacks

This topic describes the increasing occurrence of network attacks.

The figure shows how much faster DoS and worm attacks are occurring. For example, SQL Slammer was able to saturate a GigabitEthernet link within one minute. Network administrators have less and less time to react to DoS and worm attacks before their networks become unavailable. Therefore, the network has to be designed to be self-defending by using tactics link Scavenger-Class QoS.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

The Impact of an Internet Worm Attack

This topic describes the impact that a worm has on the network and network services.

The figure illustrates the impact that a worm can cause due to its speed of propagation, resulting in a DoS as it consumes network resources.

In the example, a worm is introduced into the network via a laptop PC, which in turn infects an e-mail server, which results in the e-mail server infecting other servers, while the worm is propagated to all ends of the enterprise network. Without a QoS mitigation strategy, this example is very likely to happen.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

QoS Tools and Tactics for Security

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

You can use these three QoS tools to mitigate the impact of out-of-profile traffic flows, including the mitigation of DoS attacks:

- Control Plane Policing (CPP [or CoPP for the Catalyst 6500 implementation]): The majority of traffic travels through the router via the data plane; however, a Route Processor must handle certain packets, such as routing updates, keepalives, and network management. This is often referred to as control and management plane traffic. Because the Route Processor is critical to network operations, any service disruption to the Route Processor or the control and management planes can result in business-impacting network outages. A DoS attack targeting the Route Processor, which can be perpetrated either inadvertently or maliciously, typically involves high rates of traffic that result in excessive CPU utilization on the Route Processor.

  CPP addresses the need to protect the control and management plane, ensuring routing stability, availability, and packet delivery. It uses a dedicated control-plane configuration via the Modular QoS command-line interface (CLI), or MQC, to provide filtering and rate-limiting capabilities for control plane packets. Because CPP filters traffic destined to the Route Processor, it is important to understand the legitimate traffic destined for the Route Processor prior to deployment. Configuring CPP policies without this knowledge may result in the blockage of critical traffic.

- Data plane policing: Data plane policing is the actual policing of data traffic. Understanding what is a normal profile behavior for users and servers is key to setting up policers to re-mark and drop packets. Out-of-profile behavior could be an increase in data traffic while a worm is being propagated over the enterprise network. Normal profile behavior is characterized by what is expected from end-user and server traffic.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

- NBAR known-worm policing: NBAR can detect known worms and, by marking down the traffic, drop the traffic generated by the intruder. There are only a few actual worms that can be identified by NBAR. After traffic has been generated by a known worm, it can be dropped immediately on a branch router or LAN edge.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Control Plane Policing

This topic describes CPP.

The figure shows how, to protect the Control Plane (CP) on a router from a DoS attack, the CPP feature treats the CP as a separate entity with its own ingress (input) and egress (output) ports. Because CPP treats the CP as a separate entity, a set of rules can be established and associated with the ingress and egress ports. When configured, these rules are applied after the packet has been determined to have the CP as its destination or when a packet exits from the CP. You can configure a service policy to prevent unwanted packets from progressing after a specified rate limit has been reached; for example, you can limit all SYN packets under TCP that are destined for the CP to a maximum rate of one megabit per second.

To further understand these concepts and to configure CPP, refer to *Control Plane Policing* at http://www.cisco.com/univercd/cc/td/doc/product/software/ios124/124cg/hqos_c/part20/ch05/hrtlimt.htm

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Data Plane Policing

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

Scavenger-Class QoS DoS/Worm Mitigation Strategy

- **Worms and DoS are not normal, they are anomalies.**
- **Understand the specific threat vectors so that you can mitigate anomalous behavior that could impact the network.**
- **Protect the integrity of the network.**
  - **Keep the network equipment alive and manageable.**
  - **Keep mission-critical data flowing.**
- **Implement an integrated network security architecture.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-10

Data Plane policing is the actual policing of data traffic. Understanding what is a normal profile behavior for users and servers and what is not is key in setting up policers to remark and drop packets. Out-of-profile behavior could be categorized by the increase in data traffic as a worm is being propagated over the enterprise network. Normal profile behavior could be what is expected from end-user traffic as well as server traffic.

Understanding what is a normal traffic flow within your network allows you to define a policing strategy that can mitigate any negative impact that out-of-profile flows can have on the network.

Data plane policing allows data traffic to flow, thereby providing high network availability. A key benefit to deploying data plane policing is that it augments network security so that in the event of a DoS worm attack, out-of-profile traffic can be remarked and placed into a Scavenger class, policed aggressively, and dropped.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Scavenger-Class QoS DoS/Worm Mitigation Strategy (Cont.)

- **The Scavenger class is an Internet 2 draft Specification for a "less-than-best effort" service.**
- **There is an implied "good faith" commitment for the "best effort" traffic class.**
  - **It is generally assumed that at least some network resources will be available for the default class.**
- **Scavenger class markings can be used to distinguish out-of-profile (abnormal) traffic flows from in-profile (normal) flows.**
  - **The scavenger class marking is DSCP CS1 (8).**
- **Scavenger traffic is assigned a "less-than-best effort" queuing treatment whenever congestion occurs.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-11

The Scavenger class 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 organization. Assigning Scavenger class traffic to a minimal bandwidth queue forces the traffic to reduce to almost nothing during times of congestion, but allows it to be available if bandwidth is not being used for business purposes. A Scavenger class traffic strategy allows normal traffic flows to not be impacted by spike traffic flows such as those that could be experienced in the event of a DoS worm attack.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Scavenger-Class QoS DoS/Worm Mitigation Strategy (Cont.)

- **All end systems generate traffic spikes.**
- **Sustained traffic loads beyond "normal" from each source device are considered suspect and marked as scavenger (DSCP CS1).**
- **No dropping at campus access-edge, only re-marking.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-12

The figure shows an example of what happens to out-of-profile traffic when data policing is applied.

End systems can generate traffic spikes at times, but usually those spikes are not sustained. However, you do not want a traffic spike to impact the rest of the network, so you can mark down the excessive traffic to a class that gets only about one to two percent bandwidth. Once the traffic is re-marked to scavenger class, it is policed aggressively.

The same concept applies to DoS worm attack mitigation. Worms can cause sustained spikes by effected end systems, but once this out-of-profile traffic is detected, it can be marked down to a scavenger class and policed aggressively until a security application within the network removes it.

The key is not to police the out-of-profile traffic at the access edge, but rather as the traffic enters into the core distribution layer. By using this principle, the out-of-profile traffic never negatively impacts the over-all network.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Scavenger-Class QoS DoS/Worm Mitigation Strategy (Cont.)

Queuing only engages if links become congested.

- When congestion occurs, drops will also occur.

Scavenger class QoS allows for increased intelligence in the dropping decision.

- **"Abnormal" traffic flows will be dropped aggressively.**
- **"Normal" traffic flows will continue to receive network service.**

Queuing will engage when links become congested and Traffic previously marked as scavenger is dropped aggressively.

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-14

The key point of this strategy is that legitimate traffic flows that temporarily exceed thresholds are not penalized by Scavenger-class QoS. Only sustained, abnormal streams generated simultaneously by multiple hosts (indicating DoS worm attacks) are subject to aggressive dropping—and such dropping only occurs after in-profile traffic has been fully serviced. The campus uplinks are not the only points in the network where congestion can occur. Typically, WAN and Virtual Private Network (VPN) links are the first to congest. However, Scavenger-class "less-than-best-effort" queuing should be provisioned on ***all*** network devices in a consistent manner (according to hardware capabilities).

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

NBAR Known-Worm Policing

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

With the **ip nbar custom** command, you can specify your own match criteria to identify TCP- or UDP-based applications across a range of ports, as well as on specific ports, in addition to protocols and applications identified by NBAR or via downloaded Packet Description Language Modules (PDLMs) imported to NBAR. You can specify a string or value to match at a specified byte offset within the packet payload. You can create more than 30 custom PDLMs and give them names with the **ip nbar custom** command.

Using NBAR User-Defined Application Classification enables you to specify your own criteria to match a string or numeric value inside the data packet in order to identify application traffic.

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Integrating Security Through QoS

This topic describes the impact of integrating security through QoS.

With appropriate security deployed in the network as well as the appropriate QoS tools deployed, you can mitigate excessive traffic that could literally make a network shut down. The figure depicts the minimized impact a worm attack would have if Cisco Security and QoS are in place.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

- **There is an increase in DoS attacks since the early 2000s.**
- **DoS attacks fall into two classes: Spoofing and flooding**
- **QoS tools can be used mitigate flooding attacks.**
- **Mitigation of flood attacks can come from policing traffic flows by re-marking and dropping out-of-profile network behavior.**
- **QoS tools and tactics:**
  - **Control Plane Policing (CoPP)**
  - **Data plane policing plus scavenger class**
  - **NBAR signature of known worms for policing**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-21

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Summary

This topic summarizes the key points discussed in this module.

Module Summary

- **QoS best practices include:**
  - **Ensuring that real-time traffic gets priority with minimal delay**
  - **Ensuring that mission-critical traffic is serviced appropriately**
  - **Not assigning too many applications to mission-critical traffic class**
  - **Ensuring that adequate bandwidth is provisioned for the default class**
  - **Establishing proper trust boundary; classifying and marking traffic as close to the source as possible**
  - **Using class-based policing to rate-limit certain non-business-related traffic**
- **QoS SLA is a key differentiator for service providers. SLAs typically include three to five classes. Additional classes not visible to customers may exist at the edge (for example, for management/traffic control).**
- **Different QoS design options are required for managed versus unmanaged CE.**
- **DiffServ backbone allows support multiple classes of traffic with different under/overprovisioning ratios per service class.**
- **DiffServ can reduce the bandwidth requirement on the network while achieving the same SLA when compared to the non-DiffServ case.**

© 2006 Cisco Systems, Inc. All rights reserved. QoS v2.2—9-1

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

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

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

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

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

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

References

For additional information, refer to these resources:

- For more information on QoS classification best practices, refer to *Classification & Marking* at this URL: http://www.cisco.com/pcgi-bin/Support/browse/psp_view.pl?p=Internetworking:Classification_and_Marking&viewall=true
- For more information on Enterprise QoS Design, refer to *Cisco Enterprise Quality of Service Solution Reference Network Design Guide* at this URL: http://www.cisco.com/application/pdf/en/us/guest/netsol/ns432/c649/ccmigration_09186a008049b062.pdf
- For more information on end-to-end QoS implementation, refer to *Implementing DiffServ for End-to-End Quality of Service Overview* at this URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt7/qcfdfsrv.pdf

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check

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

Q1) Which three QoS tools are used to manage the delay, delay variation (jitter), bandwidth, and packet-loss parameters on a network? (Choose three.) (Source: Understanding Traffic Classification Best Practices)

A) coding/decoding (codec) tools
B) flow-control (windowing) tools
C) scheduling tools
D) provisioning tools
E) classification tools

Q2) Which statement is true regarding the bandwidth requirement of data applications? (Source: Understanding Traffic Classification Best Practices)

A) It is not possible to provide a blanket rule for provisioning data bandwidth because the bandwidth requirements vary greatly from application to application (and even between versions of the same applications).
B) All ERP applications have similar bandwidth requirements.
C) All data applications should be provisioned a minimum bandwidth guarantee matching their data rate + 20 percent.
D) ERP applications should use the LLQ and be provisioned a maximum bandwidth guarantee matching their data rate.

Q3) Which three are QoS requirements for voice? (Choose three.) (Source: Understanding Traffic Classification Best Practices)

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

Q4) Which three statements are general guidelines for implementing campus QoS? (Choose three.) (Source: Deploying End-to-End QoS)

A) Use multiple queues on the transmit interfaces to minimize the potential for dropped or delayed traffic caused by transmit buffer congestion.
B) Classification or marking should be done at the high-speed core layer.
C) Most of the more complex QoS configurations, like traffic shaping, header compression, LFI, LLQ, and WRED, occur at the CE WAN edge router.
D) Establish a proper trust boundary. For example, at the access-layer switches, only trust the IP Phone CoS marking, but not the PC CoS marking.
E) Always use NBAR to classify traffic because all Cisco Catalyst switches support NBAR.
F) QoS in the campus is generally not a buffer management issue as much as it is a bandwidth management issue. Therefore, link efficient mechanisms should be implemented between the access and distribution layer links.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q5) The Cisco Catalyst 2950 switch has four transmit queues with a single drop threshold. By default, the transmit queues are serviced with which type of scheduling? (Source: Deploying End-to-End QoS)

A) priority
B) WRR
C) MDRR
D) FIFO

Q6) Which two behaviors are typical PHB behaviors required for the EF traffic class? (Choose two.) (Source: Deploying End-to-End QoS)

A) policed to a maximum bandwidth by LLQ
B) guaranteed a minimum bandwidth by CBWFQ
C) the dropping method should be tail-drop
D) the dropping method should be WRED to prevent congestion

Q7) Based on the following configuration, which two statements are true? (Choose two.) (Source: Deploying End-to-End QoS)

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

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

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Q8) What is wrong with the following configuration? (Source: Deploying End-to-End QoS)

```
class-map match-all PREMIUM
  match ip dscp ef
class-map match-all BUSINESS
  match ip dscp af31
!
policy-map IN-POLICY
!
class PREMIUM
  police percent 25
   exceed-action drop
!
class BUSINESS
  police percent 56
   exceed-action set-mpls-exp-transmit 3
   violate-action set-mpls-exp-transmit 2
!
 class class-default
   set mpls experimental 0
!
interface Serial0/0.1 point-to-point
 ip address 10.32.14.2 255.255.255.0
 frame-relay interface-dlci 16
 service-policy input IN-POLICY
```

Q9) If the service provider core network is using Cisco GSR as the P routers, what will be the queuing mechanism used on the Cisco GSR P routers? (Source: Deploying End-to-End QoS)

A) MDRR
B) LLQ
C) CBWFQ
D) D-WFQ
E) IP RTP Priority

Q10) The QoS requirements on the CE and PE router differ depending on which factor? (Source: Deploying End-to-End QoS)

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

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.

Module Self-Check Answer Key

Q1) C, D, E

Q2) A

Q3) A, B, C

Q4) A, C, D

Q5) A

Q6) A, C

Q7) A, F

Q8) The conform actions are missing from the configuration.

Q9) A

Q10) B

 © 2006 Cisco Systems, Inc.

The PDF files and any printed representation for this material are the property of Cisco Systems, Inc., for the sole use by Cisco employees for personal study. The files or printed representations may not be used in commercial training, and may not be distributed for purposes other than individual self-study.