End-to-End QoS Network Design
Quality of Service for Rich-Media & Cloud Networks
Second Edition

Tim Szigeti, CCIE No. 9794
Robert Barton, CCIE No. 6660
Christina Hattingh
Kenneth Briley Jr., CCIE No. 9754

Copyright © 2014 Cisco Systems, Inc.
Published by:
Cisco Press
800 East 96th Street
Indianapolis, IN 46240 USA

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the publisher, except for the inclusion of brief quotations in a review.

Printed in the United States of America
First Printing November 2013
Library of Congress Control Number: 2013950000
ISBN-10: 1-58714-369-0

Warning and Disclaimer
This book is designed to provide information about designing a network with end-to-end quality of service. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an “as is” basis. The authors, Cisco Press, and Cisco Systems, Inc., shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

Trademark Acknowledgments
All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc. cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.
Corporate and Government Sales

The publisher offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales, which may include electronic versions and/or custom covers and content particular to your business, training goals, marketing focus, and branding interests. For more information, please contact: U.S. Corporate and Government Sales 1-800-382-3419 corpsales@pearsontech-group.com

For sales outside of the U.S. please contact: International Sales international@pearsoned.com

Feedback Information

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members from the professional technical community.

Readers’ feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us through e-mail at feedback@ciscopress.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.
About the Authors

**Tim Szigeti**, CCIE No. 9794, is a senior technical leader in the Systems Design Unit at Cisco Systems, where his role is to design network architectures for enterprise mobility solutions. He has specialized in quality of service technologies for the past 15 years, during which time he has authored many technical papers, design guides, and two Cisco Press books: *End-to-End QoS Network Design* (version 1) and *Cisco TelePresence Fundamentals*.

**Robert Barton**, CCIE No. 6660, is located in Vancouver, where he lives with his wife and two children. He graduated from the University of British Columbia with a degree in engineering physics, and is a registered professional engineer. Rob holds dual CCIEs, in Routing and Switching and Security, and was also the first CCDE in Canada. Rob joined Cisco from ArrowPoint Communications, where he worked as a data center specialist supporting many of the largest corporations in Canada. In the time since ArrowPoint was acquired by Cisco, Rob has worked as a public sector systems engineer, primarily focused on wireless and security architectures. Currently, Rob is working on SmartGrid network technologies, including smart meter and intelligent substation design.

**Christina Hattingh** spent 13 years as a senior member of the technical staff in Unified Communications (UC) in the Enterprise Networking Routing Group (formerly Services Routing Technology Group or SRTG) at Cisco Systems. The SRTG products, including the Cisco 2900/3900 and 2800/3800 series ISR platforms and their predecessors, were the first Cisco platforms to converge voice, data, and video traffic and services on IP networks by offering TDM gateway interfaces, WAN interfaces, call control, and QoS features. The ISR series of routers often live at smaller remote offices and therefore at the edge of the WAN, where the need for QoS services is most sensitive. In this role, Christina spoke at Cisco Live conferences, trained Cisco sales staff and Cisco resale partners on router-based UC technologies, authored several Cisco Press books, and advised customers on UC network deployment and design, including QoS designs and helping them through the TDM to SIP trunk industry transition.

**Kenneth Briley, Jr.**, CCIE No. 9754 is a technical lead in the Network Operating Systems Technology Group at Cisco Systems. For over 10 years, he has specialized in quality of service design and implementation in customer environments, alignment of QoS features and functions, and the marketing of new products that leverage QoS technologies. During this time, he has written several deployment guides and whitepapers, presented at Cisco Live, and most recently has focused on the convergence of wired and wireless quality of service.
About the Technical Reviewers

John Johnston, previously CCIE No. 5232, is a technical marketing engineer for Cisco Systems. His focus is on mobile security technology and design validation. John has more than 19 years of experience in IP internetworking, including the design and implementation of enterprise networks. Before joining Cisco Systems, John provided network design support for Fortune 500 companies. He holds a BSEE from the UNC-Charlotte.

Roland Saville is a Technical Leader for the Systems Development Unit (SDU) at Cisco, focused on developing best-practice design guides for enterprise network deployments. He has more than 18 years of experience at Cisco as a Systems Engineer, Product Manager, Consulting Systems Engineer, Technical Marketing Engineer, and Technical Leader. During that time, he has focused on a wide range of technology areas, including the integration of voice and video onto network infrastructures, network security, wireless LAN networking, RFID, energy management, Cisco TelePresence, and BYOD. He has also spent time focusing on the retail market segment. Prior to Cisco, he spent eight years as a Communications Analyst for Chevron Corporation. Roland holds a bachelor’s of science degree in electrical engineering from the University of Idaho and an MBA degree from Santa Clara University. He co-authored the book *Cisco TelePresence Fundamentals*, is a member of the IEEE, and has 12 U.S. patents.
Dedications

Tim Szigeti:

I find myself in a dedication dilemma.

On the one hand, I already went to great lengths to explain why not dedicating the first edition of this book to my wife would have been a fatal mistake. Since then, I've gone on to dedicate my second book to my son, and now I have a beautiful daughter who deserves a dedication too (and whose arrival, incidentally, actually delayed the release of this edition by a couple of months).

So, the question becomes: Are dedications—as their definition implies—exclusive to a given book? Or can these be edition-specific? Or perhaps the more important question is: Do I really think it wise to get into a debate over semantics with my wife, who has a double-major in both English and philosophy?

So I'll play the political game and try to weakly rationalize a compromise: The first edition of this book was dedicated to Lella. The second will be to Lella 2.0, or as she's more commonly known, Isla.

Besides, I've already witnessed how much my daughter values my books. For example, over the past few months, she's had two copies of my previous book under her crib, slightly elevating one end to alleviate nighttime gas. Since she wakes up happy and smiling every morning, I'll infer from this her appreciation of the practical benefits of my work. Furthermore, she's always ready to gnaw and drool on my books until they're nice and soggy, and since pure happiness is expressed during the process, I'll attribute this to her esteem of the quality of the authorship.

And so, to my beautiful little girl, I wish to dedicate to you this work. I really don't know how I ever managed to finish it, seeing as how little you let me sleep over the past few months! I know you'll probably never read it, but that's not the point. I just want you to know you were always on my mind and made working on it virtually impossible! And I'm so very happy it's all done with now, so that I can spend more time playing with you and letting you continue wrapping me tightly around your tiny little finger!
Rob Barton:

This book is dedicated to my two wonderful boys, Adrian and Matthew. It’s not that I expect you to actually pick up the book and try to become QoS experts, or that I am even trying to encourage you toward a career in network design or engineering, although these are noble pursuits. Rather, the lesson that writing this book has reminded me of is that you only grow as a person when you recognize the space you are in and make the decision to do something new. Oftentimes, we don’t know what direction our efforts will take us in, but when you make the mindful choice to do something that is difficult, challenging, and can cause you more than a little pain along the way, you grow. No muscle ever grew without the fibers being damaged through exercise, and so is it too with all aspects of life. My hope is that this book will inspire you throughout your life to look for opportunities for growth—be it artistic, mental, professional, physical, or spiritual. This book is for you.

Christina Hattingh:

To Robert Verkroost and my parents for their unfailing encouragement and support.

Kenneth Briley, Jr.:

As this is my first book, I’d like to heed Tim’s advice and dedicate it to my beautiful wife Mirah for fear of the aforementioned transgression. To Mirah, who incidentally read and approved this dedication, and her countless hours devoted to resolving numerous grammatical errors and listening to me drone on about how incredibly interesting QoS is. To our growing family; Lukas, Erik and Max: please don’t grow up too fast, and remember that all things are possible.
Acknowledgments

Tim Szigeti:

First, I'd like to thank all the readers of the first edition who made it the success that it has become. There aren't many technology books that are still being steadily purchased nearly 10 years after their release. And a special thanks to the reviewers who have posted comments for it; I cannot express the pride and appreciation I felt when I saw five-star ratings for our book on Amazon. Thank you!

Thanks to my director, Neil Anderson, for long recognizing the critical role of QoS across all our networking systems and solutions and ensuring that it was always properly addressed. Thanks, too, to Greg Edwards in helping to define and articulate various end-to-end QoS strategies.

Thank you Fred Baker for your guidance and direction in both defining and interpreting various QoS standards. Thanks, too, to James Polk for continuing to push the envelope in the IETF to define what tomorrow's QoS models are going to look like.

I'd like to thank the Vancouver Cisco office lab administrator, Mojan Mobasser, for all her diligence in sourcing and arranging access to equipment. Similar thanks are extended to Dawid Brink for letting me use his Nexus boxes—no questions asked!

Farther east, I'd also like to extend thanks to the Toronto Bell Canada team for allowing me extended access to their ASR and CRS labs. Similar thanks, but in the opposite geographic direction, go out to Lim Fung in our Singapore office for providing me access to his labs also.

I'd like to extend sincere thanks to Tim Stevenson for his amazing technical expertise, particularly relating to data center platforms. You really helped demystify a lot of hardware architectural questions I was grappling with. Thanks, Tim!

Also I'd like to thank Lukas Krattiger in Switzerland for hours of research, testing, and correspondence to ensure that we properly wrapped our arms around Nexus 7000 QoS. Thanks for all your insight, patience, and hard work, Lukas!

Additionally, I'd like to thank Lucien Avramov for sharing his work on data center QoS and allowing me to borrow from it. Thank you too, Mike Herbert—wherever you are—for getting the ball rolling in this area.

I'd like to thank also the Cisco product teams that listened to the feedback we offered as a result of producing this book so as to continue to improve our QoS feature sets. This includes Albert Mitchell and the Catalyst 2K/3K team for implementing our latest designs into a new version of AutoQoS. Thanks also to Sarath Subrahmanya and Ramachandra Murthy in India for taking to heart our suggestions on WLC QoS feature enhancements. Kudos also go out to Oz Ben-Rephael and team in Israel for continuing to develop NBAR2 application signatures, including for our own Jabber application.

Thanks to the Cisco Press team. Brett Bartow: Thanks for taking on this project and allowing us to thoroughly update and expand on our original work in a comprehensive manner. We appreciate that you didn't blow a gasket when we exceeded our targeted
page count again, again, and again—to a final tune of target +50%! Thanks also for delaying this publication by a couple of months, letting me focus on my family as my daughter was born.

Thank you Chris Cleveland for making the review process so easy. Your comments and accommodation were very much appreciated and really helped polish this work. Thank you, too, Seth Kerney for coordinating the copy review. And also thanks to Vanessa Evans for ensuring that we always had everything we needed at every step of the way.

I'd like to extend exceptional thanks to our technical editors Roland Saville and John Johnston. Roland: You're one of the smartest persons I've had the pleasure of working with—and in so many fields. I don't know how your brain doesn't explode! You know I like to think of you as a “philosopher engineer,” because you can take almost any design recommendation and find the corner-case counterargument where it breaks down and doesn't apply. That's critically important to the process because by seeing from a distance where things can break you continually save us tremendous amounts of time in the lab, as well as ensuring the overall quality of the final designs. Thank you, too, JJ! You allowed me unfettered access to your massive labs in RTP and helped me along every step of the way. Your attention to detail is so impressive that I'm nearly spooked by your ability to catch the tiniest errors while reviewing hundreds of pages of configurations!

Finally, I owe a great deal of gratitude to my co-authors:

Ken: Thanks for your impressive knowledge and flexibility that you demonstrated by being able to jump right in and seamlessly adapt your research to our work in such an intuitive and cohesive manner. I've enjoyed working with you on many projects for the past decade and look forward to many more collaborations. Thanks again, Ken!

Christina: Thanks so much for coming out of retirement to work on one more project. Even though you're on the road more than Jack Kerouac these days, it was a real pleasure working with you again! Thanks for donning your QoS hat for us once again and bringing all your knowledge and experience to the table to help make this such a solid work.

Rob: Over the past 20 years we've been friends, classmates, roommates, workmates, “second-best” men at each other's weddings, and now co-authors. Your courage and determination are very inspiring. I honestly don't know if I would have taken my CCIE if I hadn't watched you do it. Same goes with running half-marathons (and one-day marathons!) Thanks for all your tremendous work on this project. It certainly was not for the faint-hearted, as every time we turned around we seemed to uncover yet another rabbit hole of technical issues that required yet more research and testing to be done. Thanks for sticking with it and seeing it through, Rob. But then again, that's just the kind of friend you are.

Rob Barton:

To begin, I would like to thank my very forgiving colleagues in the Cisco Vancouver office who have suffered through two years of trying to depend on an attention divided systems engineer who was more interested in solving theoretical QoS problems
than in helping his customers. Special thanks to my Cisco account team partner, Mike MacDonald, for his long-suffering patience, my manager, Ralph Wright, who enthusiastically supported this effort and always offered many words of encouragement, and to my director, Bill Kastelic, who eagerly gave me the flexibility to do this project. None of this would have been possible without the support from you guys.

I would also like to thank my lab administrator, Mojan Mobasser, for helping to get lab gear when I needed it the most. Testing these QoS solutions involved a lot of lab time, and without your support we would not have been able to build and test these solutions.

Special thanks goes out to Ian Procyk and my co-author Ken Briley who helped test some of the more difficult wireless scenarios. As well, I would like to thank Larry Ross for the many hours of emails and phone conversations discussing various wireless QoS solutions with me. Also thanks goes out to Kangwarn Chinthammit for helping with the AVC section review, and Scott Wainer who helped with the GET VPN work. All you guys were like my technical conscious during this project.

I’d also like to thank Bruno Wollmann from Terra-Comm Solutions who, while discussing my presentation at Cisco Live last year, introduced me to the concept of combining DMVPN with GET VPN to solve a real-world performance issue related to VoIP, which I think has made a great addition to the GET VPN chapter.

Chris Cleveland and Brett Bartow, thanks so much for your hard work on this project and supporting us all the way through. This project turned into a much bigger undertaking than any of us had expected, and instead of trying to apply your own QoS mechanism on our output, you let the creative juices flow, and in the end helped support a substantial work of technical literature.

Lastly, I’d like to thank Tim Szigeti. Not only have you been one of my closest friends for more than 20 years, you are also an inspiring engineer. Yes, I said engineer, the word you always tease me with. I can clearly remember the day this project started two and a half years ago; we were rewarding ourselves with a well-earned breakfast at the White Spot after one of our half-marathon training runs. I was complaining that your first edition of the End-to-End QoS book, while being a great book, was hopelessly out of date. Your response to me was unforgettable: “So why don’t you help me write a new one?” That day was the start of this project, and although it was a long and difficult undertaking, it has also been an immensely rewarding experience. Thanks, Tim!

Kenneth Briley, Jr.:

First off I’d like to thank Roland Saville, for his guidance and clever insight when we worked through QoS on the Converged Access platforms.

To Stephen Orr, wireless is now awesome, before it was an illusion – thanks for the brilliant and oh so colorful commentary.

Many thanks to Tripti Agarwal, Saravanam Radhakrishnan, Anuj Kumar, and Bhavana Kudipudi without that team we would have never been able to deliver such a versatile platform.
## Contents at a Glance

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>xxxvi</td>
</tr>
<tr>
<td><strong>Part I: QoS Design Overview</strong></td>
<td></td>
</tr>
<tr>
<td>Chapter 1 Introduction and Brief History of QoS and QoE</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2 IOS-Based QoS Architectural Framework and Syntax Structure</td>
<td>13</td>
</tr>
<tr>
<td>Chapter 3 Classification and Marking</td>
<td>31</td>
</tr>
<tr>
<td>Chapter 4 Policing, Shaping, and Markdown Tools</td>
<td>59</td>
</tr>
<tr>
<td>Chapter 5 Congestion Management and Avoidance Tools</td>
<td>83</td>
</tr>
<tr>
<td>Chapter 6 Bandwidth Reservation Tools</td>
<td>99</td>
</tr>
<tr>
<td>Chapter 7 QoS in IPv6 Networks</td>
<td>111</td>
</tr>
<tr>
<td>Chapter 8 Medianet</td>
<td>117</td>
</tr>
<tr>
<td>Chapter 9 Application Visibility Control (AVC)</td>
<td>135</td>
</tr>
<tr>
<td><strong>Part II: QoS Design Strategies</strong></td>
<td></td>
</tr>
<tr>
<td>Chapter 10 Business and Application QoS Requirements</td>
<td>163</td>
</tr>
<tr>
<td>Chapter 11 QoS Design Principles and Strategies</td>
<td>189</td>
</tr>
<tr>
<td>Chapter 12 Strategic QoS Design Case Study</td>
<td>215</td>
</tr>
<tr>
<td><strong>Part III: Campus QoS Design</strong></td>
<td></td>
</tr>
<tr>
<td>Chapter 13 Campus QoS Design Considerations and Recommendations</td>
<td>223</td>
</tr>
<tr>
<td>Chapter 14 Campus Access (Cisco Catalyst 3750) QoS Design</td>
<td>247</td>
</tr>
<tr>
<td>Chapter 15 Campus Distribution (Cisco Catalyst 4500) QoS Design</td>
<td>275</td>
</tr>
<tr>
<td>Chapter 16 Campus Core (Cisco Catalyst 6500) QoS Design</td>
<td>305</td>
</tr>
<tr>
<td>Chapter 17 Campus QoS Design Case Study</td>
<td>347</td>
</tr>
<tr>
<td><strong>Part IV: Wireless LAN QoS Design</strong></td>
<td></td>
</tr>
<tr>
<td>Chapter 18 Wireless LAN QoS Considerations and Recommendations</td>
<td>373</td>
</tr>
<tr>
<td>Chapter 19 Centralized (Cisco 5500 Wireless LAN Controller) QoS Design</td>
<td>397</td>
</tr>
<tr>
<td>Chapter 20 Converged Access (Cisco Catalyst 3850 and the Cisco 5760 Wireless LAN Controller) QoS Design</td>
<td>435</td>
</tr>
<tr>
<td>Chapter 21 Converged Access QoS Design Case Study</td>
<td>477</td>
</tr>
</tbody>
</table>
Part V: Data Center QoS Design

Chapter 22  Data Center QoS Design Considerations and Recommendations  499
Chapter 23  Data Center Virtual Access (Nexus 1000V) QoS Design  535
Chapter 24  Data Center Access/Aggregation (Nexus 5500/2000) QoS Design  561
Chapter 25  Data Center Core (Nexus 7000) QoS Design  599
Chapter 26  Data Center QoS Design Case Study  651

Part VI: WAN and Branch QoS Design

Chapter 27  WAN and Branch QoS Design Considerations and Recommendations  675
Chapter 28  WAN Aggregator (Cisco ASR 1000) QoS Design  697
Chapter 29  Branch Router (Cisco ISR G2) QoS Design  735
Chapter 30  WAN and Branch QoS Design Case Study  759

Part VII: MPLS VPN QoS Design

Chapter 31  MPLS VPN QoS Design Considerations and Recommendations  771
Chapter 32  Enterprise Customer Edge (Cisco ASR 1000 and ISR G2) QoS Design  793
Chapter 33  Service Provider Edge (Cisco ASR 9000) QoS Design  809
Chapter 34  Service Provider Core (Cisco CRS) QoS Design  845
Chapter 35  MPLS VPN QoS Design Case Study  861

Part VIII: IPsec QoS Design

Chapter 36  IPsec VPN QoS Considerations and Recommendations  871
Chapter 37  DMVPN QoS Design  893
Chapter 38  GET VPN QoS Design  921
Chapter 39  Home Office VPN QoS Case Study  943

Index  953

Part XI: Appendixes (Online)

Appendix A  AutoQoS for Medianet
Appendix B  Control Plane Policing
Contents

Introduction xxxvi

Part I: QoS Design Overview

Chapter 1  Introduction and Brief History of QoS and QoE  1

History and Evolution  2

Then  3

Now  3

Evolution of QoS  4

QoS Basics and Concepts  5

User Expectations: QoS, QoE, and QoX  5

QoS Models: IntServ and DiffServ  6

Fundamental QoS Concepts and Toolset  7

Packet Headers  8

Simplifying QoS  9

Standardization and Consistency  9

Summary  11

Further Reading  11

General  11

IntServ  12

DiffServ  12

Chapter 2  IOS-Based QoS Architectural Framework and Syntax Structure  13

QoS Deployment Principles  13

QoS Architectural Framework  14

QoS Behavioral Model  15

QoS Feature Sequencing  15

Modular QoS Command-Line Framework  16

MQC Syntax  17

Default Behaviors  19

Traffic Classification (Class Maps)  19

Definition of Policies (Policy Maps)  20

Attaching Policies to Traffic Flows (Service Policy)  22

Hierarchical QoS and HQF  23

Legacy QoS CLI No Longer Used  25

AutoQoS  26

Summary  29
Further Reading  29  
  General  29  
  AutoQoS  29  

**Chapter 3  Classification and Marking  31**  
Classification and Marking Topics  31  
  Classification and Marking Terminology  32  
Security and QoS  33  
  *Trust Boundaries*  33  
  Network Attacks  34  
Classification Challenges of Video and Wireless Traffic  34  
Marking Fields in Different Technologies  35  
*Field Values and Interpretation*  35  
Ethernet 802.1Q/p  37  
Ethernet 802.11 WiFi  38  
ATM and FR  38  
IPv4 and IPv6  39  
L2 and L3 Tunnels  39  
CAPWAP  40  
MPLS  41  
Mapping QoS Markings  41  
*Mapping L2 to L3 Markings*  41  
*Mapping Cisco to RFC 4594 Markings*  42  
*Mapping Markings for Wireless Networks*  43  
Classification Tools  44  
  *Class-Based Classification (Class Maps)*  45  
Network-Based Application Recognition  47  
*NBAR Protocols*  48  
RTP Traffic  49  
*Performance Routing*  49  
*Metadata Classification*  50  
Marking Tools  50  
  *Class-Based Marking (Class Maps)*  50  
*Effects of Feature Sequence*  52  
*Mapping Markings with the Table Map Feature*  52  
Marking (or Re-Marking) with Policing  53  
AutoQoS Marking  54
Chapter 4 Policing, Shaping, and Markdown Tools 59

Policing and Shaping Topics 59
- Policing and Shaping Terminology 60
  - Placing Policers and Shapers in the Network 61
  - Tail Drop and Random Drop 61
  - Re-Mark/Markdown 62
  - Traffic Types to Police and Shape 62
- Token Bucket Algorithms 62
  - Types of Policers 64
    - Single-Rate Two-Color Policers 64
      - RFC 2697 Single-Rate Three-Color Policers 65
      - RFC 2698 Dual-Rate Three-Color Policers 66
- Security and QoS 68
  - Policing Tools 68
    - Policers as Markers 68
      - Class-Based Policing (Policy Maps) 69
      - Multi-Action Policing 70
    - Hierarchical Policing 71
      - Percentage-Based Policing 72
        - Color-Aware Policing 73
      - Policing as Part of Low-Latency Queuing 73
    - Control Plane Policing 74
      - Unconditional Packet Drop 75
- Traffic Shaping Tools 75
  - Class-Based Shaping (Policy Maps) 76
    - Hierarchical Class-Based Shaping 77
      - Percentage-Based Shaping 77
  - Legacy Shaping Tools 78
ATM Traffic Shaping  78
Frame Relay Traffic Shaping  78
Recommendations and Guidelines  79
Summary  80
Further Reading  80
  General  80
  DiffServ Policing Standards  80
  Policing  80
  Shaping  81

Chapter 5  Congestion Management and Avoidance Tools  83
Congestion Management and Avoidance Topics  84
  Congestion Management and Avoidance Terminology  84
  Congestion Management and Congestion Avoidance  85
Scheduling Algorithms  85
Levels of Queuing  85
Queuing and Scheduling Tools  86
  Class-Based Queuing (Policy Maps)  86
  Class-Based Weighted Fair Queuing  88
  Low-Latency Queuing  88
  Queuing Below Layer 3: Tx-Ring Operation  91
Congestion Avoidance Tools  92
  Random Early Detection  93
  Weighted Random Early Detection  93
Recommendations and Guidelines  95
Summary  96
Further Reading  96
  Queuing  96
  Congestion Avoidance  96

Chapter 6  Bandwidth Reservation Tools  99
Admission Control Tools  100
Resource Reservation Protocol  101
  RSVP Overview  101
  RSVP Proxy  102
  RSVP Deployment Models  103
  Basic RSVP Design (IntServ/DiffServ Model)  104
  Advanced RSVP Design (IntServ/DiffServ Model)  105
Chapter 7  QoS in IPv6 Networks  111
IPv6 and QoS Overview  111
QoS Tools for IPv6  112
  QoS Feature Support for IPv6  112
  Packet Headers, Classification, and Marking  112
    Packet Classification  113
    Packet Marking  114
  Policing and Shaping  115
Recommendations and Guidelines  115
Summary  116
Further Reading  116

Chapter 8  Medianet  117
An Introduction to Medianet  117
Medianet Architecture and Framework  119
Medianet Features and Capabilities  120
  Autoconfiguration  121
  Auto Smartports  121
  AutoQoS  121
  Media Monitoring  122
  Mediatrace  122
  Performance Monitor  125
  IPSLA Video Operation (Traffic Simulator, IPSLA VO)  127
  Media Awareness  128
  Flow Metadata  129
  Network Based Application Recognition  130
  Media Services Interface  132
  Media Services Proxy  132
Summary  133
Further Reading  133
  Overviews  133
Chapter 9  Application Visibility Control (AVC)  135
AVC Use Cases  136
How AVC Works  138
The AVC Building Blocks  140
Building Block 1: NBAR2  140
NBAR2 Protocol Discovery  142
NBAR2 MQC Traffic Classification  144
Building Block 2: Flexible NetFlow  147
Flexible NetFlow Key Fields and Non-Key Fields  148
Configuration of FNF  149
Building Block 3: AVC Management and Reporting  152
Insight Reporter  153
Building Block 4: AVC QoS Controls  154
Deploying AVC QoS Controls at the WAN Edge  154
Deploying AVC QoS Controls at the Internet Edge  156
Performance Considerations When Using AVC  159
Summary  160
Additional Reading  161

Part II: QoS Design Strategies
Chapter 10  Business and Application QoS Requirements  163
Global Trends in Networking  164
The Evolution of Video Applications  164
The Explosion of Media  166
The Phenomena of Social Networking  167
The Bring Your Own Device Demand  167
The Emergence of Bottom-Up Applications  168
The Convergence of Media Subcomponents Within Multimedia Applications  168
The Transition to High-Definition Media  169
QoS Requirements and Recommendations by Application Class  169
Voice  170
Video Applications  171
Chapter 11  QoS Design Principles and Strategies  189

QoS Best-Practice Design Principles  189
  Hardware Versus Software QoS Best Practices  190
  Classification and Marking Best Practices  191
  Policing and Markdown Best Practices  192
  Queuing and Dropping Best Practices  192
  EF Queue Recommendations: The 33% LLQ Rule  193
  AF Queue Recommendations  195
  DF Queue Recommendations  195
  Scavenger Class Queue Recommendations  195
  WRED Recommendations  197

QoS Design Strategies  198
  Four-Class Model QoS Strategy  198
  Eight-Class Model QoS Strategy  200
Chapter 12 Strategic QoS Design Case Study 215
Tifosi Software Inc.: Company Overview 215
Original (Four-Class) QoS Model 215
Business Catalysts for QoS Reengineering 216
Proposed (Eight-Class) QoS Model 217
“Layer 8” Challenges 219
Summary 221
Additional Reading 221

Part III: Campus QoS Design
Chapter 13 Campus QoS Design Considerations and Recommendations 223
MLS Versus MQC 225
Default QoS 226
Internal DSCP 226
Trust States and Operations 227
Trust Boundaries 230
DSCP Transparency 231
Port-Based QoS Versus VLAN-Based QoS Versus Per-Port/Per-VLAN QoS 232
EtherChannel QoS 234
Campus QoS Models 235
Ingress QoS Models 235
Egress QoS Models 238
Campus Port QoS Roles 239
Campus AutoQoS 241
Control Plane Policing 243
Summary 244
Additional Reading 246
Chapter 14 Campus Access (Cisco Catalyst 3750) QoS Design 247
Cisco Catalyst 3750 QoS Architecture 248
QoS Design Steps 249
  Enabling QoS 250
  Ingress QoS Models 250
  Trust Models 251
Classification and Marking Models 254
Classification, Marking, and Policing Models 256
Queuing Models 260
  Ingress Queuing Model 261
  Egress Queuing Models 265
Additional Platform-Specific QoS Design Options 271
  Per-VLAN QoS Design 271
  Per-Port/Per-VLAN QoS 272
  EtherChannel QoS Design 273
  AutoQoS SRND4 273
  Control Plane Policing 274
Summary 274
Additional Reading 274

Chapter 15 Campus Distribution (Cisco Catalyst 4500) QoS Design 275
Cisco Catalyst 4500 QoS Architecture 276
QoS Design Steps 277
Queuing Models 277
  Four-Class Egress Queuing Model 278
  Eight-Class Egress Queuing Model 281
  Twelve-Class Egress Queuing Model 284
Additional Platform-Specific QoS Design Options 289
  Access-Edge Design Options 290
  Conditional Trust Model 290
  Medianet Metadata Classification Model 292
Classification and Marking Models 293
Classification, Marking, and Policing Model 294
Per-VLAN QoS Design 297
Per-Port/Per-VLAN QoS 298
EtherChannel QoS Design 299
Flow-Based QoS 301
Chapter 16  Campus Core (Cisco Catalyst 6500) QoS Design  305

Cisco Catalyst 6500 QoS Architecture  306
QoS Design Steps  308
Queuing Models  308
Four-Class (4Q4T Ingress and 1P3Q4T Egress) Queuing Models  311
Eight-Class (8Q4T Ingress and 1P7Q4T Egress) Queuing Models  314
Twelve-Class (8Q4T Ingress and 1P7Q4T Egress) Queuing Models  318
2P6Q4T Ingress and Egress Queuing Models  328
Additional Platform-Specific QoS Design Options  329
Access-Edge Design Options  330
Conditional Trust Model  330
Classification and Marking Models  332
Classification, Marking, and Policing Model  335
Microflow Policing  341
Per-VLAN QoS Design  342
EtherChannel QoS Design  343
AutoQoS SRND4  344
Control Plane Policing  344
Summary  344
Further Reading  345

Chapter 17  Campus QoS Design Case Study  347

Tifosi Campus Access QoS Design  350
Policy 1: Access-Edge Design for Printer Endpoints (No Trust)  351
Policy 2: Access-Edge Design for Wireless Access Endpoints (DSCP Trust)  351
Policy 3: Access-Edge Design for Cisco TelePresence Endpoints (Conditional Trust)  352
Policy 4: Access-Edge Design for Cisco IP Phones or PCs (Conditional Trust and Classification and Marking)  352
Eight-Class 1P1Q3T Ingress Queuing Design  355
Eight-Class 1P3Q3T Egress Queuing Design  357
Policy 5: Access Layer Uplink Design  359
Tifosi Campus Distribution QoS Design  360
   Policy 6: Distribution Layer Downlink Ports (Catalyst 4500E Supervisor 7-E)  360
   Policy 7: Distribution Layer Distribution-Link / Core-Uplink Ports  362
Tifosi Campus Core QoS Design  364
   Policy 8: Core Layer (10GE) Downlink Design  364
   Policy 9: Core Layer (40GE) Core-Link Design  368
Summary  370
Further Reading  371

Part IV: Wireless LAN QoS Design

Chapter 18  Wireless LAN QoS Considerations and Recommendations  373
Comparing QoS in Wired and Wireless LAN Environments  374
WLAN QoS Building Blocks  376
   The Distributed Coordination Function  376
   CSMA/CA  377
   The DCF Contention Window  378
IEEE 802.11e and Wireless Multimedia (WMM)  382
   Retrofitting DCF: Enhanced Distributed Channel Access  382
   Access Categories  383
   Arbitration Interframe Spacing  385
   Contention Window Enhancements  386
   Transmission Opportunity  388
   802.11e TSpec: Call Admission Control  388
QoS Design Considerations  389
   Defining Upstream and Downstream Traffic Flow  389
   QoS Mapping and Marking Considerations  390
   The Upstream QoS Marking Strategy  392
   The Downstream QoS Marking Strategy  394
Summary  395
Additional Reading  396

Chapter 19  Centralized (Cisco 5500 Wireless LAN Controller) QoS Design  397
QoS Enforcement Points in the WLAN  398
Managing QoS Profiles in the Wireless LAN Controller  399
   QoS Marking and Conditional Trust Boundaries  399
   WLAN QoS Profiles  400
Chapter 21  Converged Access QoS Design Case Study  477

Tifosi Converged Access QoS Design: Wired  481

Policy 1: Access-Edge Design for Wired Printer Endpoints (No Trust)  481
Policy 2: Access-Edge Design for Wired Access Endpoints (DSCP Trust)  481
Policy 3: Access-Edge Design for Cisco TelePresence Endpoints (Conditional Trust)  482
Policy 4: Access-Edge Design for Cisco IP Phones and PCs (Conditional Trust and Classification and Marking)  482
Policy 5: Access-Edge Wired Queuing Design  485

Tifosi Converged Access QoS Design: Wireless  488

Policy 6: Access-Edge Design for Mobile Wireless Clients (Dynamic Policy with and Classification & Marking)  489
Policy 7: Access-Edge Wireless Queuing Design  491
Policy 8: SSID Bandwidth Wireless Bandwidth Allocation Between Guest and Enterprise SSIDs (SSID Policy to Separate Bandwidth Distribution)  492
Policy 9: CT 5760 Wireless LAN Controller Uplink Ports  493

Cisco Identity Services Engine  495
Summary  496
Additional Reading  496

Part V: Data Center QoS Design

Chapter 22  Data Center QoS Design Considerations and Recommendations  499

Data Center Architectures  500

High-Performance Trading Data Center Architectures  500
Big Data (HPC/HTC/Grid) Architectures  501
Virtualized Multiservice Data Center Architectures  503
Secure Multitenant Data Center Architectures  505
Massively Scalable Data Center Architectures  506

Data Center QoS Tools  507

Data Center Bridging Toolset  508
Ethernet Flow Control: IEEE 802.3x  508
Priority Flow Control: IEEE 802.1Qbb  510
Skid Buffers and Virtual Output Queuing  512
Enhanced Transmission Selection: IEEE 802.1Qaz  514
Congestion Notification: IEEE 802.1Qau  515
Data Center Bridging Exchange: IEEE 802.1Qaz + 802.1AB  516
QoS Design Steps  569
Ingress QoS Models  569
Trust Models  570
  Trusted Server Model  570
  Untrusted Server Model  570
Classification and Marking Models  572
  Single-Application Server Model  573
  Multi-Application Server Model  576
Application Policing Server Model  578
  Modifying the Ingress Buffer Size  580
Egress Queuing Models  582
  Four-Class Model  582
  Eight-Class Model  587
Additional QoS Designs Options  592
  Nexus 5500 L3 QoS Configuration  592
  Nexus 2000 Fabric Extender QoS  593
  Using the network-qos Policy to Set MTU  597
Summary  597

Additional Reading  598

Chapter 25  Data Center Core (Nexus 7000) QoS Design  599

Nexus 7000 Overview  600
Nexus 7000 M2 Modules: Architecture and QoS Design  604
  M2 QoS Design Steps  607
  M2 Queuing Models  607
  M2 Default Queuing Models  608
  M2 Four-Class (4Q2T Ingress / 1P3Q4T Egress) Queuing Model  610
  M2 Eight-Class (8Q2T Ingress / 1P3Q4T Egress) Queuing Model  615
  M2 OTV Edge Device QoS Design  621
Nexus 7000 F2 Modules: Architecture and QoS Design  623
  F2 QoS Design Steps  625
  F2 Network QoS Policy Design  625
  F2 Queuing Models  630
  F2 Default Queuing Models  631
  F2 Four-Class (4Q1T Ingress / 1P3Q1T Egress) Queuing Model  634
  F2 Eight-Class (4Q1T Ingress / 1P3Q1T Egress) Queuing Model  634
  FEX QoS Design  638
Additional M2/F2 QoS Design Options 638
   Trusted Server Model 638
   Untrusted Server Model 638
   Single-Application Server Marking Model 642
   Multi-Application Server Classification and Marking Model 642
   Server Policing Model 643
   DSCP-Mutation Model 645
CoPP Design 648
Summary 648
Further Reading 649

Chapter 26 Data Center QoS Design Case Study 651
Tifosi Data Center Virtual Access Layer Nexus 1000V QoS Design 655
   Policy 1: Trusted Virtual Machines 655
   Policy 2: Single-Application Virtual Machine 655
   Policy 3: Multi-Application Virtual Machine 656
   Policy 4: Network-Edge Queuing 657
Tifosi Data Center Access/Aggregation Layer Nexus 5500/2000 QoS Design 659
   Policy 5: Trusted Server 660
   Policy 6: Single-Application Server 660
   Policy 7: Multi-Application Server 661
   Policy 8: Network-Edge Queuing Policy 662
Tifosi Data Center Core Layer Nexus 7000 QoS Design 666
   Policy 9: Network-Edge Queuing (F2 Modules) 666
   Policy 10: Network-Edge Queuing (M2 Modules) 668
   Policy 11: DSCP Mutation for Signaling Traffic Between Campus and Data Center 671
Summary 672
Further Reading 673

Part VI: WAN and Branch QoS Design

Chapter 27 WAN and Branch QoS Design Considerations and Recommendations 675
WAN and Branch Architectures 677
Hardware Versus IOS Software QoS 678
Latency and Jitter 679
Tx-Ring 682
Chapter 29  Branch Router (Cisco ISR G2) QoS Design  735
Cisco ISR G2 QoS Architecture  736
QoS Design Steps  738
Ingress QoS Models  738
  Medianet Classification Models  738
  Medianet Application-Based Classification and Marking Model  739
  Medianet Application-Group-Based Classification Model  743
  Medianet Attribute-Based Classification Model  744
  NBAR2 Classification Models  744
  NBAR2 Application-Based Classification and Marking Model  745
  NBAR2 Application-Group-Based Classification Model  748
  NBAR2 Attribute-Based Classification Model  748
  Custom-Protocol NBAR2 Classification  752
Egress QoS Models  753
  Four-Class Model  754
  Eight-Class Model  754
  Twelve-Class Model  754
Additional Platform-Specific QoS Design Options  757
  RSVP  757
  AutoQoS SRND4  757
  Control Plane Policing  757
Summary  757
Further Reading  758

Chapter 30  WAN and Branch QoS Design Case Study  759
Policy 1: Internal (PLIM) QoS for ASR 1000  761
  Policy 1a: SIP-Based PLIM QoS  762
  Policy 1b: SPA-Based PLIM QoS  762
Policy 2: LAN-Edge QoS Policies  763
Policy 3: WAN Edge QoS Policies  765
Summary  768
Further Reading  769

Part VII: MPLS VPN QoS Design

Chapter 31  MPLS VPN QoS Design Considerations and Recommendations  771
MPLS VPN Architectures  772
MAN and WAN Ethernet Service Evolution  773
Sub-Line-Rate Ethernet Design Implications  775
QoS Paradigm Shift  779
Service Provider Class of Service Models  781
MPLS DiffServ Tunneling Modes  781
  Uniform Mode  782
  Short Pipe Mode  783
  Pipe Mode  784
Enterprise-to-Service Provider Mapping  785
  Mapping Real-Time Voice and Video  785
  Mapping Control and Signaling Traffic  786
  Separating TCP from UDP  786
  Re-Marking and Restoring Markings  787
MPLS VPN QoS Roles  787
Summary  789
Further Reading  790

Chapter 32  Enterprise Customer Edge (Cisco ASR 1000 and ISR G2) QoS Design  793
QoS Design Steps  794
Ingress QoS Models  795
Egress QoS Models  795
  Sub-Line-Rate Ethernet: Hierarchical Shaping and Queuing Models  795
  Known SP Policing Bc  796
  Unknown SP Policing Bc  797
Enterprise-to-Service Provider Mapping Models  798
  Four-Class Enterprise Model Mapped to a Four-CoS Service Provider Model  798
  Eight-Class Enterprise Model Mapped to a Six-CoS Service Provider Model  800
  Twelve-Class Enterprise Model Mapped to an Eight Class-of-Service Service Provider Model  803
Summary  808
Further Reading  808

Chapter 33  Service Provider Edge (Cisco ASR 9000) QoS Design  809
QoS Architecture  810
QoS Design Steps  814
MPLS DiffServ Tunneling Models  814
  Uniform Mode MPLS DiffServ Tunneling  815
  Uniform Mode Ingress Policer  816
Chapter 34  Service Provider Core (Cisco CRS) QoS Design 845
QoS Architecture 846
QoS Design Steps 849
SP Core Class-of-Service QoS Models 849
  Four-Class-of-Service SP Model 850
  Four-Class-of-Service Fabric QoS Policy 850
  Four-Class-of-Service Interface QoS Policy 853
Six-Class-of-Service SP Core Model 854
  Six-Class-of-Service Fabric QoS Policy 855
  Six-Class-of-Service Interface QoS Policy 856
Eight-Class-of-Service SP Core Model 857
  Eight-Class-of-Service Fabric QoS Policy 857
  Eight-Class-of-Service Interface QoS Policy 858
Summary 860
Additional Reading 860

Chapter 35  MPLS VPN QoS Design Case Study 861
Policy 1: CE Router Internal QoS (Cisco ASR 1000) 863
Policy 2: CE Router LAN-Edge QoS Policies 863
Policy 3: CE Router VPN-Edge QoS Policies 863
Policy 4: PE Router Internal QoS (Cisco ASR 9000) 866
Policy 5: PE Router Customer-Edge QoS 866
Policy 6: PE Router Core-Edge QoS 867
Policy 7: P Router Internal QoS (Cisco CRS-3) 868
Policy 8: P Router Interface QoS 868
Summary 868
Additional Reading 868

Part VIII: IPsec QoS Design

Chapter 36 IPsec VPN QoS Considerations and Recommendations 871
IPsec VPN Topologies 871
Standard IPsec VPNs 872
Tunnel Mode 872
Transport Mode 873
IPsec with GRE 873
Remote-Access VPNs 874
QoS Classification of IPsec Packets 875
The IOS Preclassify Feature 877
MTU Considerations 880
How GRE Handles MTU Issues 881
How IPsec Handles MTU Issues 881
Using the TCP Adjust-MSS Feature 883
Compression Strategies Over VPN 885
TCP Optimization Using WAAS 885
Using Voice Codecs over a VPN Connection 886
cRTP and IPsec Incompatibilities 887
Antireplay Implications 888
Summary 891
Additional Reading 891

Chapter 37 DMVPN QoS Design 893
The Role of QoS in a DMVPN Network 895
DMVPN Building Blocks 895
How QoS Is Implemented in a DMVPN? 895
DMVPN QoS Configuration 896
Next-Hop Routing Protocol 897
The Need for a Different Approach to QoS in DMVPNs 898
The Per-Tunnel QoS for DMVPN Feature 899
DMVPN QoS Design Example 900
DMVPN QoS Design Steps 902
Configuring the Hub Router for Per-Tunnel QoS 902
Part XI: Appendixes (Online)

Appendix A   AutoQoS for Medianet
Appendix B   Control Plane Policing
“Aren’t we done with QoS yet?”

That’s a question I get from time-to-time, which I like to answer along the lines of “As soon as we’re done with availability and security, we’ll be done with QoS also.”

What I’m trying to express—although cheekily—is that although QoS has been around for a while, it is a foundational network infrastructure technology (the same as high-availability technologies and security technologies). And these foundational technologies will always prove to be integral components of any networking system, being present at the platform level, at the place in-the network (PIN) level and ultimately at the end-to-end network level.

Furthermore, such foundational network technologies are constantly evolving and expanding to meet new business and technical requirements. Such has been the case with QoS since the first edition of this work was published nearly 10 years ago.

For example, consider just one QoS-dependent application: video.

In 2004, there were really only two flavors of video traversing most enterprise networks: streaming video (unidirectional flows that benefited from both network- and application-level buffering to offset variations in transmission delays) and video conferencing (bidirectional 384-Kbps or 768-Kbps streams between dedicated hardware-based systems). So, we went into our massive Cisco Validation Labs in Research Triangle Park in North Carolina and hammered out best-practice designs to accommodate these two categories of video. We were done, right?

Wrong.

In the years that followed, codec and hardware advances made video production more cost-efficient and accessible, such that today nearly everyone with a smartphone has the ability to shoot high-definition video anytime and anywhere. Similarly, with the advent of social networking websites, video sharing and distribution suddenly became possible by anyone, anywhere (and that on a global scale!). Finally, video consumption also became possible anytime, anywhere, and on any device—thanks to advances in hardware and in wireless networking technologies.

That being the case, video is now the most dominant type of network traffic on the Internet and is expected to reach 90 percent within in a few years. Furthermore, there are many new forms and variations of video traffic, such as TelePresence, IP video surveillance, desktop video, and digital signage (just to name a few). And each of these types of video has unique service level requirements that must be met to ensure a high quality of experience by the end user. And thus, we circle back to QoS, which represents the enabling technologies to provide this quality of experience.

And that’s just one application.

Advances in areas of data center and cloud networking, in addition to wireless networking, all have had corresponding impacts on QoS network designs.
Hence, a new edition of this book.

Another reason behind this second edition is to reflect the evolution of industry standards relating to QoS. Cisco has long advocated following industry standards and recommendations whenever deploying QoS, because this simplifies QoS designs, extends QoS policies beyond an administrative domain, and improves QoS policy effectiveness between administrative domains. Therefore, new standards, RFCs, and proposals have had—and will continue to have—a major impact on current and future strategic QoS designs.

A third key reason behind this new edition is that every network platform detailed in the original book has been replaced or significantly upgraded. So, the latest platforms (at the time of this writing) have been featured in this second version, with over a dozen Cisco product families being represented. In fact, nearly every design chapter features a different Cisco platform that suits the role being discussed, whether the role is a data center virtual switch, a branch router, a wireless LAN controller, a campus distribution switch, a WAN aggregator, a service provider core router, or so on.

And finally, QoS is a comprehensive and complex subject, one that entails a significant amount of fundamental technological concepts as well as platform-specific implementation detail. Therefore, it is often valuable for network administrators to have a single common reference on the subject, such as this book, which overviews all the relevant tools, presents various end-to-end strategies, and details platform-specific design recommendations for every major shipping Cisco platform.

And no, we’re not done with QoS yet!

**Objectives of This Book**

The main objective of this book is to present—in a comprehensive and cohesive manner—the many aspects of quality of service design, including an overview of the tools, strategic and tactical design recommendations, and platform-specific configuration details. Therefore, novice to advanced network administrators alike can benefit from this volume as a single handy reference on this topic.

In addition, this exercise has produced multiple platform-specific configurations that can be viewed as QoS templates. As such, these templates can be considered roughly 80 percent of a generic enterprise or service provider QoS solution (borrowing from Pareto’s 80/20 rule), to which another 20 percent of customizing and tailoring can be done to reach a final customer-specific solution. Considerations and rationales behind the presented designs are all explained so that administrators are fully informed of the rationale behind the designs and therefore can confidently modify these to meet their own specific requirements and constraints.

A key approach that we’ve used throughout this configuration-rich book is to incorporate inline explanations of configurations. In this way, QoS-relevant commands are highlighted and detailed line-by-line to explicate the function of each element and clarify how these parts make up the solution as a whole.
To complement these line-by-line design recommendations, related verification commands are also incorporated. These verification commands are presented in context with the design examples, with specific details of what-to-look-for being highlighted and explained. These verification examples are therefore significantly richer in relevance than most such examples presented in hardware/software documentation, and they allow network administrators to confirm quickly whether the recommended designs have been deployed correctly.

Finally, each design section has a case study chapter at the end that ties together many of the strategic principles, tactical recommendations, and platform-specific considerations that have been presented within the section. These case studies illustrate how to take generic and abstract design concepts and mold them to meet specific customer requirements. These case studies are indicative of what can be expected in real-life production environments. Each of these case study examples spans multiple devices, thus highlighting critical interrelationships. Furthermore, all case study chapters form respective parts of a single integrated end-to-end QoS network design.

Who Should Read This Book?

The primary reader of this book is the network administrator tasked with deploying QoS technologies. By extension, this group may also include other related IT professionals, such as systems administrators, audio/video specialists, VoIP specialists, and operations staff.

In addition, some readers may include technical decision makers tasked with evaluating the strategy and feasibility of QoS deployments, in addition to the drafting of implementation plans and phases toward these goals.

Yet another group of readers includes system engineers, partners, trainers, and other networking professionals who need to ramp-up technically on QoS technologies and designs, both for practical deployment purposes and to achieve various Cisco certifications.

Prerequisites are minimal, as the opening section of this book covers QoS technologies in high-to-mid-level technical detail, including protocols, tools, and relevant standards. In addition, each chapter includes extensive references for Additional Reading for more detailed information for readers unfamiliar with specific concepts discussed.

Because the content of the book ranges from a high level to a very low level of technical detail, it is suitable for a wide range of audiences, from intermediate to expert.

How This Book Is Organized

This book is organized into 39 chapters distributed across 8 parts, and includes 2 appendices. Although this book can be read cover to cover, this organization allows readers to easily identify chapters of direct interest, thus facilitating the use of this book as a handy reference work. The eight parts of this book are described below:
Part I, “QoS Design Overview,” introduces readers to QoS technologies, presenting a brief history and an architectural framework for these tools. Following this, groups of QoS tools are overviewed, including classification and marking tools, policing and shaping tools, queuing and dropping tools, bandwidth-reservation tools, and advanced tools like Medianet and application visibility and control.

Part II, “QoS Design Strategies,” breaks away from a purely technical discussion to take a higher-level view of how business requirements drive QoS design. Application service-level requirements are analyzed, as are strategic QoS design best practices. This section concludes with the first case study chapter, illustrating the considerations that factor into defining an end-to-end QoS design strategy.

Part III, “Campus QoS Design,” begins the exercise applying strategic QoS models to a tactical place in the network (PIN), which in this case is the enterprise campus. Campus-specific design considerations and recommendations are discussed at length, and subsequent chapters specialize in design recommendations for the access, distribution, and core layers of the campus network. A campus QoS design case study chapter completes the section.

Part IV, “Wireless LAN QoS Design,” applies the strategic QoS models to the enterprise wireless LAN. Because WiFi is a unique media, as compared to the rest of the network, additional concepts need to be covered to explain how QoS can be achieved over-the-air. These considerations include the introduction of the Enhanced Distributed Coordination Function as well as IEEE 802.11e/Wireless Multimedia QoS. Following this, QoS design chapters address both the centralized wireless LAN controller deployment model and the new wired-and-wireless converged access deployment model. The section finishes with a WLAN QoS design case study.

Part V, “Data Center QoS Design,” continues the application of QoS strategies, but this time to the data center network. Because of the convergence of storage-area networks and local-area networks within the data center, certain protocols require a completely lossless service that traditional QoS tools cannot guarantee. Therefore, data center-specific QoS tools are discussed, including the data center bridging toolset, which can be leveraged to guarantee such a lossless service. Following this, QoS design chapters address the virtual access layer, access and aggregation layers, and the core layer of data center networks. This part closes with a data center QoS design case study.

Part VI, “WAN and Branch QoS Design,” expands the scope of discussion beyond the local area and applies strategic QoS principles to the wide-area network. QoS designs are presented for both WAN aggregation routers and for branch routers. This part ends with a WAN QoS design case study.

Part VII, “MPLS VPN QoS Design,” continues the wide-area discussion but addresses QoS strategies for MPLS VPN networks, taking the perspectives of both the enterprise customer and the service provider into account in the end-to-end design. Design chapters are presented for the enterprise customer-edge router, the provider-edge router and the provider core routers. This section finishes with a case study on MPLS VPN QoS design.
Part VIII, “IPsec QoS Design,” concludes the discussion by applying strategic QoS principles to IPsec VPNs. QoS designs are detailed for both Dynamic Multipoint VPNs and Group Encrypted Transport VPNs.

An overview on each of the 39 chapters (and the 2 appendixes) follows.

- **Chapter 1, “Introduction and Brief History of QoS and QoE”:** Provides a brief history lesson on quality of service and quality of experience evolution, introducing fundamental QoS concepts, standards, and the evolutionary changes necessitating a second edition of this book.

- **Chapter 2, “IOS-Based QoS Architectural Framework and Syntax Structure”:** Overviews how QoS tools interrelate, and introduces Cisco’s IOS-based Modular QoS command-line interface (MQC), the common syntax structure for configuring QoS across most Cisco platforms.

- **Chapter 3, “Classification and Marking Tools”:** Describes the various classification options for distinguishing one packet from another, which is the requisite first step in providing differentiated services. Also discussed are various marking options so that packets do not have to be reclassified at every network node.

- **Chapter 4, “Policing, Shaping, and Markdown Tools”:** Discusses various tools that can be used to meter and regulate packet flows, including policers (which drop excess traffic), shapers (which delay excess traffic) and markers (which re-mark excess traffic).

- **Chapter 5, “Congestion Management and Avoidance Tools”:** Considers options on how to deal with bottlenecks in the network, by addressing both queuing tools (to determine which packets get priority or preferential treatment during congestion), and early-dropping tools (to reduce the probability of congestion).

- **Chapter 6, “Bandwidth-Reservation Tools”:** Introduces the concepts of bandwidth reservations and endpoint/infrastructure signaling to communicate how and when such reservations are to be made.

- **Chapter 7, “QoS in IPv6 Networks”:** Examines IPv6 packet formats, classification and marking options, and how QoS tools are to be configured in IPv6 networks or in mixed IPv4 and IPv6 networks.

- **Chapter 8, “Medianet”:** Gives a brief overview of the Medianet architecture, with particular focus on the aspects of Medianet specific to QoS configuration and monitoring.

- **Chapter 9, “Application Visibility and Control”:** Presents deep packet inspection technologies for application identification, classification, and monitoring and how these can be used within the network.

- **Chapter 10, “Business and Application QoS Requirements”:** Examines current business trends impacting QoS designs and various application-class QoS requirements.
Chapter 11, “QoS Design Principles and Strategies”: Combines the QoS tools and business requirements presented in preceding chapters and formulates these into QoS strategic models to address basic, intermediate, and advanced requirements.

Chapter 12, “Strategic QoS Design Case Study”: This first case study in the series introduces a fictional company, Tifosi Software, and discusses the business and technical considerations that come into play when defining an end-to-end QoS strategy.

Chapter 13, “Campus QoS Design Considerations and Recommendations”: Overviews various considerations and recommendations relating to campus QoS design, including trust boundaries, per-port versus per-VLAN design options, and EtherChannel QoS considerations.

Chapter 14, “Campus Access (Cisco Catalyst 3750) QoS Design”: This first platform-specific design chapter details best practice QoS designs at a configuration level for Cisco Catalyst 3750 series switches in the role of a campus access layer edge switch.

Chapter 15, “Campus Distribution (Cisco Catalyst 4500) QoS Design”: This design chapter details configuration recommendations for a Cisco Catalyst 4500 series switch in the role of a campus distribution layer switch. Additional designs include details on how this switch can be configured as a campus access-edge switch also.

Chapter 16, “Campus Core (Cisco Catalyst 6500) QoS Design”: This design chapter details configuration recommendations for a Cisco Catalyst 6500 series switch in the role of a campus core layer switch. Additional designs include details on how this switch can be configured as a campus access-edge or distribution layer switch as well.

Chapter 17, “Campus QoS Design Case Study”: This case study chapter describes how Tifosi Software has applied their strategic QoS design model to their campus network consisting of Cisco Catalyst 3750, 4500 and 6500 series switches.

Chapter 18, “Wireless LAN QoS Considerations and Recommendations”: Overviews various considerations and recommendations relating to wireless LAN QoS design and introduces WLAN QoS tools such as the Enhanced Distributed Coordination Function and Wireless Multimedia QoS.

Chapter 19, “Centralized (Cisco 5500 Wireless LAN Controller) QoS Design”: This design chapter details both GUI and CLI configuration recommendations for centralized wireless LAN controller (WLC) deployment models, featuring the Cisco 5500 WLC.

Chapter 20, “Converged Access (Cisco Catalyst 3850 and the Cisco 5760 Wireless LAN Controller QoS Design”: This design chapter details configuration recommendations for converged access WLAN deployment models, featuring the Cisco Catalyst 3850 series switch and the Cisco 5760 WLC.
- Chapter 21, “Converged Access QoS Design Case Study”: This case study chapter describes how Tifosi Software has applied their strategic QoS design model to their wired-and-wireless converged access LAN network consisting of Cisco Catalyst 3850 series switches and the Cisco 5760 WLC.

- Chapter 22, “Data Center QoS Design Considerations and Recommendations”: Overviews various considerations and recommendations relating to data center QoS design and introduces the data center bridging toolset.

- Chapter 23, “Data Center Virtual Access (Nexus 1000V) QoS Design”: This design chapter details configuration recommendations for a Cisco Nexus 1000V series virtual switch in the role of a data center access layer switch.

- Chapter 24, “Data Center Access/Aggregation (Nexus 5500/2000) QoS Design”: This design chapter details configuration recommendations for a Cisco Nexus 5500 series switch, which may include Cisco Nexus 2000 series Fabric Extenders, in the role of a data center access/aggregation switch.

- Chapter 25, “Data Center Core (Nexus 7000) QoS Design”: This design chapter details configuration recommendations for a Cisco Nexus 7000 series switch in the role of a data center core switch. QoS designs for both M-Series and F-Series modules are detailed.

- Chapter 26, “Data Center QoS Design Case Study”: This case study chapter describes how Tifosi Software has applied their strategic QoS design model to their data center network, consisting of Cisco Nexus 1000V, 5500/2000 and 7000 series switches.

- Chapter 27, “WAN and Branch QoS Design Considerations and Recommendations”: Overviews various considerations and recommendations relating to WAN QoS design, including hardware versus software considerations, latency and jitter targets, and bandwidth-reservation options.

- Chapter 28, “WAN Aggregator (Cisco ASR 1000) QoS Design”: This design chapter details configuration recommendations for a Cisco ASR 1000 series router in the role of a WAN aggregation router. WAN media featured includes leased lines, ATM, and Packet-Over-SONET.

- Chapter 29, “Branch Router (Cisco ISR G2) QoS Design”: This design chapter details configuration recommendations for a Cisco ISR G2 series router in the role of a branch router, featuring Medianet and AVC designs.

- Chapter 30, “WAN and Branch QoS Design Case Study”: This case study chapter describes how Tifosi Software has applied their strategic QoS design model to their wide-area network, consisting of Cisco ASR 1000 and ISR G2 series routers.

- Chapter 31, “MPLS VPN QoS Design Considerations and Recommendations”: Overviews various considerations and recommendations relating to MPLS VPN QoS design, both from an enterprise and from a service provider perspective, including enterprise-to-provider mapping models and MPLS DiffServ tunneling modes. In
addition, this design section features carrier Ethernet as a WAN media.

- **Chapter 32, “Enterprise Customer Edge (Cisco ASR 1000 and ISR G2) QoS Design”**: This design chapter details configuration recommendations for a Cisco ASR 1000 or ISR G2 series router in the role of an enterprise customer-edge router interfacing with a MPLS VPN service provider.

- **Chapter 33, “Service Provider Edge (Cisco ASR 9000) QoS Design”**: This design chapter details configuration recommendations for a Cisco ASR 9000 series router in the role of a service provider edge router.

- **Chapter 34, “Service Provider Core (Cisco CRS) QoS Design”**: This design chapter details configuration recommendations for a Cisco CRS-3 series router in the role of a service provider core router.

- **Chapter 35, “MPLS VPN QoS Design Case Study”**: This case study chapter describes how Tifosi Software has adapted their strategic eight-class enterprise QoS model to integrate with their service provider's six class-of-service model, featuring Cisco ISR G2, ASR 1000, ASR 9000, and CRS-3 series routers.

- **Chapter 36, “IPsec VPN QoS Considerations and Recommendations”**: Overviews various considerations and recommendations relating to IPsec VPN QoS design, including classification of encrypted packets, MTU considerations, and anti-replay implications.

- **Chapter 37, “DMVPN QoS Design”**: This design chapter details configuration recommendations for Cisco ASR 1000 and ISR G2 routers in the roles of DMVPN hub-and-spoke routers (respectively).

- **Chapter 38, “GET VPN QoS Design”**: This design chapter details configuration recommendations for Cisco ISR G2 routers in the roles of GET VPN routers.

- **Chapter 39, “Home Office VPN QoS Case Study”**: This case study chapter describes how Tifosi Software has adapted their strategic QoS model over a DMVPN to provide telecommuting services to employees in their home offices. This case study features Cisco ASR 1002 series routers at the headend and ISR 881 series routers connected behind a broadband modem via Ethernet at the home office.

- **Appendix A, “AutoQoS for Medianet”**: This online appendix overviews the latest evolution of the AutoQoS feature, which is based on the same QoS designs presented in this book. Detailed syntax is presented for the first platforms to support this feature, including the Cisco Catalyst 3750 and 4500 series switches.

- **Appendix B, “Control Plane Policing”**: This online appendix overviews the control plane policing feature, which applies a QoS function (of policing) to a virtual interface (the control plane) to harden the network infrastructure from denial-of-service or worm attacks. Best-practice recommendations and configurations are presented for this feature.
This page intentionally left blank
The primary role of quality of service (QoS) in the campus distribution switch is to manage packet loss. Therefore, the distribution switch should trust differentiated services code point (DSCP) markings on ingress (as these have been previously set by access-edge switches) and perform both ingress (if required and supported) and egress queuing, as illustrated in Figure 15-1.

![Figure 15-1: Campus Distribution Switch Port QoS Roles](image)

The Cisco Catalyst 4500E Supervisor 7-E is a platform well suited to the role of a campus distribution switch and therefore is featured in this design chapter.

Incidentally, the QoS design requirements of a Catalyst 4500E Supervisor 7-E in the role of a distribution switch are generally equivalent to the requirements of a campus core switch.
Cisco Catalyst 4500 QoS Architecture

From a QoS perspective, the Cisco Catalyst 4500-E Supervisor 7-E is nearly identical to the Supervisor 6-E platform and the Catalyst 4500-X, because all of these platforms are Modular QoS command-line interface (MQC) based. However, earlier Catalyst 4500 platforms (such as the Supervisor II-Plus through Supervisor V-10GE) are Multi-Layer Switch (MLS)-QoS-based platforms and are referred to as Classic Supervisors.

Note  QoS design for these older Classic Supervisors is beyond the scope for this design chapter. However, you can find design guidance for these platforms at http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND_40/QoSCampus_40.html#wp1099634.

Figure 15-2 illustrates the QoS architecture for this Catalyst 4500E Supervisor 7-E (hereafter referred to simply as the Catalyst 4500) platform.

**Figure 15-2  Cisco Catalyst 4500 QoS Architectural Model**

QoS is enabled by default on all MQC-based platforms, which includes the Catalyst 4500. In addition, by default, all ports are set to a trust-DSCP/trust-CoS state.

In the MQC-based Catalyst 4500, QoS policies are applied as follows:

1. The incoming packet is classified (based on different packet fields, receive port, or VLAN) to belong to a traffic class.

2. Depending on the traffic class and configured polices, the packet is policed, which may result in the packet being dropped or re-marked.
Chapter 15: Campus Distribution (Cisco Catalyst 4500) QoS Design

3. After the packet has been marked/re-marked, it is looked up for forwarding. This action obtains the transmit port and VLAN to transmit the packet.

4. The packet is classified in the output direction based on the transmit port or VLAN/marking.

5. Depending on the output policies, the packet is policed, and may be dropped or re-marked.

6. The transmit queue for the packet is determined based on the traffic class and the configured egress queuing policies.

7. The transmit queue state is dynamically monitored via Dynamic Buffer Limiting (DBL) and drop threshold configuration to determine whether the packet should be dropped or queued for transmission.

8. If eligible for transmission, the packet is assigned to a transmit queue.

Based on these QoS operations, the design steps for configuring QoS on the Catalyst 4500 in the role of a distribution switch are discussed next.

QoS Design Steps

While there are two explicit QoS policy requirements of a distribution switch (namely to trust DSCP on ingress and queuing policies), because of the default QoS settings on MQC-based platforms there is effectively only a single step to configuring QoS on a Catalyst 4500 in this role:

1. Configure the ingress QoS model—which is recommended to be DSCP trust (and which is enabled by default on all MQC-based platforms).

Note This step may include ingress queuing policies on platforms which support this feature (however, the Catalyst 4500 does not support ingress queuing).

2. Configure egress queuing.

Queuing Models

Ingress queuing is not supported on the Catalyst 4500; only egress queuing is supported.

Note Other ingress QoS policies (including trust, classification, marking, and policing) are all supported; only ingress queuing is not supported on this platform.
The Catalyst 4500 supports a strict-priority hardware queue with (up to) seven additional nonpriority hardware queues. In addition, the Catalyst 4500 supports DSCP-to-queue mapping.

At the time of this writing, DSCP-based weighted random early detection (WRED) is not supported on the Catalyst 4500 platform. However, the Catalyst 4500 family uses a platform-specific congestion avoidance algorithm to provide active queue management (AQM), namely Dynamic Buffer Limiting (DBL). DBL tracks the queue length for each traffic flow in the switch. When the queue length of a flow exceeds its limit, DBL drop packets or sets the Explicit Congestion Notification (ECN) bits in the packet headers. The DBL algorithm can identify belligerent flows (that is, unchecked/nonadaptive/inelastic flows) and drop these more aggressively. Belligerent flows can use excessive bandwidth and switch buffers, resulting in poor application performance for well-behaved flows. Therefore, DBL can induce not only random “probabilistic drops” (in a manner similar to WRED), but also “belligerent flow drops,” both of which are counted and displayed via the show policy-map interface command output on classes where DBL has been enabled (as demonstrated later in Example 15-4).

Therefore, the egress queuing model for the Catalyst 4500 platform can be expressed as 1P7Q1T+DBL.

**Note** DBL is unique to the Catalyst 4500 platforms. At the time of this writing, there are no tuning options for DBL.

The Catalyst 4500 can be configured to support 4-class, 8-class, or 12-class queuing models, as discussed in the following sections.

**Four-Class Egress Queuing Model**

In the four-class model (illustrated in Figures 11-3 and 11-4 in Chapter 11, “QoS Design Principles and Strategies”), the application class to queue mappings are as follows:

- Real-time traffic (marked EF) is assigned to the priority queue (which may be optionally policed to 30 percent bandwidth).
- Control traffic (marked CS3) is assigned to a dedicated nonpriority queue with a 10 percent bandwidth allocation.
- Transactional data (marked AF2) is assigned to another dedicated nonpriority queue with a 35 percent bandwidth allocation with DBL enabled.
- Best-effort traffic (marked DF) is assigned to a default queue with 25 percent bandwidth allocation with DBL enabled.
**Note**  DBL is enabled only on the transactional data queue and the default queue (because real-time traffic and control traffic should never be early dropped).

**Note**  When the priority queue is configured on one class of a policy map without a policer, only bandwidth remaining percent is accepted on other classes (guaranteeing a minimum bandwidth for other classes from the remaining bandwidth of what is left after using the priority queue). However, when the priority queue is configured with a policer, either bandwidth percent or bandwidth remaining percent is accepted on the other queuing classes.

**Note**  If queuing policies are to be applied to EtherChannel interfaces, it is recommended not to police the priority queue. This is because two policy maps would be needed in this case: One policy map would be needed to police the priority queue (which would have to be applied to the logical EtherChannel interface in the egress direction), and a second policy map would be needed to define the queuing policy (using bandwidth remaining percent), which would be applied to all EtherChannel physical port-member interfaces in the egress direction. Therefore, to simplify the queuing policy and to increase its portability and modularity, the priority queue is not policed in the queuing design examples in this chapter (which necessitates the use of bandwidth remaining percent on nonpriority queues).

**Note**  Although it is true that there will be fractional differences in bandwidth allotments to an application class depending on whether bandwidth percent or bandwidth remaining percent is used. However, because these differences are relatively minor, the same numeric values are used in these examples for the sake of consistency.

Figure 15-3 illustrates the resulting four-class (1P3Q1T+DBL) egress queuing model for the Catalyst 4500.

Example 15-1 shows the corresponding configuration for four-class (1P3Q1T+DBL) egress queuing on the Catalyst 4500.
Example 15-1  Four-Class (1P3Q1T+DBL) Egress Queuing Configuration Example on a Catalyst 4500

! This section configures the class maps for the egress queuing policy
C4500(config)# class-map match-all PRIORITY-QUEUE
C4500(config-cmap)# match dscp ef
  ! VoIP (EF) is mapped to the PQ
C4500(config)# class-map match-all CONTROL-QUEUE
C4500(config-cmap)# match dscp cs3
  ! Signaling (CS3) is mapped to a dedicated queue
C4500(config)# class-map match-all TRANSACTIONAL-DATA-QUEUE
C4500(config-cmap)# match dscp af21 af22 af23
  ! Transactional Data (AF2) is mapped to a dedicated queue

! This section configures the four-class egress queuing policy map
C4500(config)# policy-map 1P3Q1T
C4500(config-pmap-c)# class PRIORITY-QUEUE
C4500(config-pmap-c)# priority
  ! Enables the priority queue
C4500(config-pmap-c)# class CONTROL-QUEUE
C4500(config-pmap-c)# bandwidth remaining percent 10
! Defines the control queue with 10% BW remaining
C4500(config-pmap-c)# class TRANSACTIONAL-DATA-QUEUE
C4500(config-pmap-c)# bandwidth remaining percent 35
dbl
! Defines a transactional data queue with 35% BW remaining + DBL
C4500(config-pmap-c)# class class-default
C4500(config-pmap-c)# bandwidth remaining percent 25
dbl
! Provisions the default/Best Effort queue with 25% BW remaining + DBL

! This section attaches the egress queuing policy to the interface(s)
C4500(config)# interface range TenGigabitEthernet 1/1-2
C4500(config-if-range)# service-policy output 1P3Q1T

**Note**  Class maps defined for egress-queuing policies require unique names from any ingress-policy class maps; otherwise, classification errors can occur due to overlapping classification logic.

You can verify the configuration in Example 15-1 with the following commands:

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

**Eight-Class Egress Queuing Model**

In the eight-class model (illustrated in Figures 11-5 and 11-6), the application class to queue mappings are as follows:

- Real-time traffic (marked EF) is assigned to the priority queue (which may be optionally policed to 10 percent bandwidth).
- Interactive video (marked AF4) is assigned to a dedicated nonpriority queue with a 23 percent bandwidth allocation with DBL enabled.
- Streaming video (marked AF3) is assigned to a dedicated nonpriority queue with a 10 percent bandwidth allocation with DBL enabled.
- Network control traffic (marked CS6) is assigned to a dedicated nonpriority queue with a 5 percent bandwidth allocation.
- Signaling traffic (marked CS3) is assigned to a dedicated nonpriority queue with a 2 percent bandwidth allocation.
- Transactional data (marked AF2) is assigned to dedicated nonpriority queue with a 24 percent bandwidth allocation with DBL enabled.
- Scavenger traffic (marked CS1) is constrained within a dedicated nonpriority queue with a 1 percent bandwidth allocation.
- Best-effort traffic (marked DF) is assigned to a default queue with 25 percent bandwidth allocation with DBL enabled.

**Note** As before, DBL is not enabled on the real-time or control traffic classes (because real-time traffic and control traffic should never be early dropped); nor would DBL be required on the scavenger class, because traffic in this class has no “good-faith” guarantee of service to begin with. Enabling DBL on the Interactive Video and Streaming Video classes assumes that the video codecs used for these flows are adaptive/elastic and therefore will adjust transmission rates in the event of congestion.

Figure 15-4 illustrates the resulting eight-class (1P7Q1T+DBL) egress queuing model for the Catalyst 4500.

**Figure 15-4** Catalyst 4500 Eight-Class (1P7Q1T+DBL) Egress Queuing Model

Example 15-2 shows the corresponding configuration for eight-class (1P7Q1T+DBL) egress queuing on the Catalyst 4500.
**Example 15-2**  *Eight-Class (1P7Q1T+DBL) Egress Queuing Configuration Example on a Catalyst 4500*

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>class-map match-all PRIORITY-QUEUE</code></td>
<td>Defines a priority queue</td>
</tr>
<tr>
<td><code>match dscp ef</code></td>
<td>VoIP (EF) is mapped to the PQ</td>
</tr>
<tr>
<td><code>class-map match-all INTERACTIVE-VIDEO-QUEUE</code></td>
<td>Interactive-Video (AF4) is assigned a dedicated queue</td>
</tr>
<tr>
<td><code>match dscp af41 af42 af43</code></td>
<td></td>
</tr>
<tr>
<td><code>class-map match-all STREAMING-VIDEO-QUEUE</code></td>
<td>Streaming-Video (AF3) is assigned a dedicated queue</td>
</tr>
<tr>
<td><code>match dscp af31 af32 af33</code></td>
<td></td>
</tr>
<tr>
<td><code>class-map match-all CONTROL-QUEUE</code></td>
<td>Network Control (CS6) is mapped to a dedicated queue</td>
</tr>
<tr>
<td><code>match dscp cs6</code></td>
<td></td>
</tr>
<tr>
<td><code>class-map match-all SIGNALING-QUEUE</code></td>
<td>Signaling (CS3) is mapped to a dedicated queue</td>
</tr>
<tr>
<td><code>match dscp cs3</code></td>
<td></td>
</tr>
<tr>
<td><code>class-map match-all TRANSACTIONAL-DATA-QUEUE</code></td>
<td>Transactional Data (AF2) is assigned a dedicated queue</td>
</tr>
<tr>
<td><code>match dscp af21 af22 af23</code></td>
<td></td>
</tr>
<tr>
<td><code>class-map match-all SCAVENGER-QUEUE</code></td>
<td>Scavenger (CS1) is assigned a dedicated queue</td>
</tr>
<tr>
<td><code>match dscp cs1</code></td>
<td></td>
</tr>
<tr>
<td><code>policy-map 1P7Q1T</code></td>
<td>This section configures the 1P7Q1T+DBL egress queuing policy map</td>
</tr>
<tr>
<td><code>class PRIORITY-QUEUE</code></td>
<td>Defines a priority queue</td>
</tr>
<tr>
<td><code>priority</code></td>
<td></td>
</tr>
<tr>
<td><code>class INTERACTIVE-VIDEO-QUEUE</code></td>
<td>Defines an interactive-video queue with 23% BW remaining + DBL</td>
</tr>
<tr>
<td><code>bandwidth remaining percent 23</code></td>
<td></td>
</tr>
<tr>
<td><code>=dbl</code></td>
<td></td>
</tr>
<tr>
<td><code>class STREAMING-VIDEO-QUEUE</code></td>
<td>Defines a streaming-video queue with 10% BW remaining + DBL</td>
</tr>
<tr>
<td><code>bandwidth remaining percent 10</code></td>
<td></td>
</tr>
<tr>
<td><code>=dbl</code></td>
<td></td>
</tr>
<tr>
<td><code>class CONTROL-QUEUE</code></td>
<td>Defines a control/management queue with 5% BW remaining</td>
</tr>
<tr>
<td><code>bandwidth remaining percent 5</code></td>
<td></td>
</tr>
<tr>
<td><code>class SIGNALING-QUEUE</code></td>
<td>Defines a signaling queue with 2% BW remaining</td>
</tr>
<tr>
<td><code>bandwidth remaining percent 2</code></td>
<td></td>
</tr>
</tbody>
</table>
You can verify the configuration in Example 15-2 with the following commands:

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

Twelve-Class Egress Queuing Model

In the 12-class model (illustrated in Figures 11-7 and 11-8), the application class to queue mappings are as follows:

- Voice (marked EF), broadcast video (marked CS5), and real-time interactive traffic (marked CS4) is all assigned to the priority queue (which may be optionally policed to 30 percent bandwidth).
- Multimedia-conferencing traffic (marked AF4) is assigned to a dedicated nonpriority queue with a 10 percent bandwidth allocation with DBL enabled.
- Multimedia-streaming traffic (marked AF3) is assigned to a dedicated nonpriority queue with a 10 percent bandwidth allocation with DBL enabled.
- Network control traffic (marked CS6), signaling traffic (marked CS3) and network management traffic (marked CS2) is all assigned to a dedicated nonpriority queue with a 10 percent bandwidth allocation; optionally, CS7 traffic may also be mapped to this queue.
- Transactional data traffic (marked AF2) is assigned to dedicated nonpriority queue with a 10 percent bandwidth allocation with DBL enabled.
- Bulk data traffic (marked AF1) is assigned to a dedicated nonpriority queue with 4 percent bandwidth allocation with DBL enabled.
- Scavenger traffic (marked CS1) is constrained within a dedicated nonpriority queue with a 1 percent bandwidth allocation.
- Best-effort traffic (marked DF) is assigned to a default queue with 25 percent bandwidth allocation with DBL enabled.

Figure 15-5 illustrates the resulting 12-class (1P7Q1T+DBL) egress queuing model for the Catalyst 4500.

![Catalyst 4500 12-Class (1P7Q1T+DBL) Egress Queuing Model]

Example 15-3 shows the corresponding configuration for 12-class (1P7Q1T+DBL) egress queuing on the Catalyst 4500.

Example 15-3  **Twelve-Class (1P7Q1T+DBL) Egress Queuing Configuration Example on a Catalyst 4500**

```plaintext
! This section configures the class maps for the egress queuing policy
C4500(config)# class-map match-any PRIORITY-QUEUE
C4500(config-cmap)# match dscp ef
C4500(config-cmap)# match dscp cs5
C4500(config-cmap)# match dscp cs4
! VoIP (EF), Broadcast Video (CS5) and Realtime Interactive (CS4)
! are all mapped to the PQ
```
End-to-End QoS Network Design

```conf
C4500(config)# class-map match-any CONTROL-MGMT-QUEUE
C4500(config-cmap)# match dscp cs7
C4500(config-cmap)# match dscp cs6
C4500(config-cmap)# match dscp cs3
C4500(config-cmap)# match dscp cs2
! Network Control (CS7), Internetwork Control (CS6),
! Signaling (CS3) and Management (CS2) are mapped
! to a Control/Management Queue
C4500(config)# class-map match-all MULTIMEDIA-CONFERENCING-QUEUE
C4500(config-cmap)# match dscp af41 af42 af43
! Multimedia Conferencing (AF4) is assigned a dedicated queue
C4500(config)# class-map match-all MULTIMEDIA-STREAMING-QUEUE
C4500(config-cmap)# match dscp af31 af32 af33
! Multimedia Streaming (AF3) is assigned a dedicated queue
C4500(config)# class-map match-all TRANSACTIONAL-DATA-QUEUE
C4500(config-cmap)# match dscp af21 af22 af23
! Transactional Data (AF2) is assigned a dedicated queue
C4500(config)# class-map match-all BULK-DATA-QUEUE
C4500(config-cmap)# match dscp af11 af12 af13
! Bulk Data (AF1) is assigned a dedicated queue
C4500(config)# class-map match-all SCAVENGER-QUEUE
C4500(config-cmap)# match dscp cs1
! Scavenger (CS1) is assigned a dedicated queue

! This section configures the 1P7Q1T+DBL egress queuing policy map
C4500(config)# policy-map 1P7Q1T
C4500(config-pmap-c)# class PRIORITY-QUEUE
C4500(config-pmap-c)# priority
! Defines a priority queue
C4500(config-pmap-c)# class CONTROL-MGMT-QUEUE
C4500(config-pmap-c)# bandwidth remaining percent 10
! Defines a control/management queue with 10% BW remaining
C4500(config-pmap-c)# class MULTIMEDIA-CONFERENCING-QUEUE
C4500(config-pmap-c)# bandwidth remaining percent 10
C4500(config-pmap-c)# dbl
! Defines a multimedia conferencing queue with 10% BW remaining + DBL
C4500(config-pmap-c)# class MULTIMEDIA-STREAMING-QUEUE
C4500(config-pmap-c)# bandwidth remaining percent 10
C4500(config-pmap-c)# dbl
! Defines a multimedia streaming queue with 10% BW remaining + DBL
C4500(config-pmap-c)# class TRANSACTIONAL-DATA-QUEUE
C4500(config-pmap-c)# bandwidth remaining percent 10
C4500(config-pmap-c)# dbl
```
Chapter 15: Campus Distribution (Cisco Catalyst 4500) QoS Design  287

! Defines a transactional data queue with 10% BW remaining + DBL
C4500(config-pmap-c)# class BULK-DATA-QUEUE
C4500(config-pmap-c)# bandwidth remaining percent 4
C4500(config-pmap-c)# dbl
! Defines a bulk data queue with 10% BW remaining + DBL
C4500(config-pmap-c)# class SCAVENGER-QUEUE
C4500(config-pmap-c)# bandwidth remaining percent 1
! Defines a (minimal) scavenger queue with 1% BW remaining/limit
C4500(config-pmap-c)# class class-default
C4500(config-pmap-c)# bandwidth remaining percent 25
C4500(config-pmap-c)#dbl
! Provisions the default/Best Effort queue with 25% BW remaining + DBL

! This section attaches the egress queuing policy to the interface(s)
C4500(config)# interface range TenGigabitEthernet 1/1-2
C4500(config-if-range)# service-policy output 1P7Q1T

You can verify the configuration in Example 15-3 with the following commands:

- show class-map
- show policy-map
- show policy-map interface (as shown in Example 15-4)

**Example 15-4  Verifying Queuing Policies on a Catalyst 4500: show policy-map interface**

C4500# show policy-map interface TenGigabitEthernet 1/1
TenGigabitEthernet1/1
Service-policy output: 1P7Q1T
Class-map: PRIORITY-QUEUE (match-any)
  102598 packets
  Match: dscp ef (46)
  102598 packets
  Match: dscp cs5 (40)
  0 packets
  Match: dscp cs4 (32)
  0 packets
  priority queue:
    Transmit: 22782306 Bytes, Queue Full Drops: 0 Packets

Class-map: CONTROL-MGMT-QUEUE (match-any)
  24847 packets
Match: dscp cs7 (56)
  0 packets
Match: dscp cs6 (48)
  0 packets
Match: dscp cs3 (24)
  24847 packets
Match: dscp cs2 (16)
  0 packets
bandwidth remaining 10 (%)
  Transmit: 24909844 Bytes, Queue Full Drops: 0 Packets

Class-map: MULTIMEDIA-CONFERENCING-QUEUE (match-all)
  22280511 packets
Match: dscp af41 (34) af42 (36) af43 (38)
bandwidth remaining 10 (%)
  Transmit: 4002626800 Bytes, Queue Full Drops: 0 Packets
  dbl
  Probabilistic Drops: 0 Packets
  Belligerent Flow Drops: 0 Packets

Class-map: MULTIMEDIA-STREAMING-QUEUE (match-all)
  0 packets
Match: dscp af31 (26) af32 (28) af33 (30)
bandwidth remaining 10 (%)
  Transmit: 0 Bytes, Queue Full Drops: 0 Packets
  dbl
  Probabilistic Drops: 0 Packets
  Belligerent Flow Drops: 0 Packets

Class-map: TRANSACTIONAL-DATA-QUEUE (match-all)
  235852 packets
Match: dscp af21 (18) af22 (20) af23 (22)
bandwidth remaining 10 (%)
  Transmit: 247591260 Bytes, Queue Full Drops: 0 Packets
  dbl
  Probabilistic Drops: 0 Packets
  Belligerent Flow Drops: 0 Packets

Class-map: BULK-DATA-QUEUE (match-all)
  2359020 packets
Match: dscp af11 (10) af12 (12) af13 (14)
Example 15-4 shows various queuing classes and their associated packet and byte counts, including 26,268 queuing drops noted on the scavenger queue.

**Additional Platform-Specific QoS Design Options**

These designs represent a generic building block for Catalyst 4500 QoS in a campus distribution switch role, but they are by no means the only design options available to you. Additional options and considerations include the following:

- Access-edge design options
- Per-VLAN QoS design
- Per-port/per-VLAN QoS design
- EtherChannel QoS design
- AutoQoS SRND4
- Control plane policing

Each of these additional QoS design options is discussed in turn.
Access-Edge Design Options

This chapter has focused on QoS designs for the Catalyst 4500 in the role of a campus distribution switch (which are generally equivalent to the QoS designs required were it serving in the role of a campus core switch). However, the Catalyst 4500 can also be deployed as a campus access switch. Therefore, a few additional design options would apply in such a role, including the following access-edge models:

- Conditional Trust Model
- Classification and Marking Model
- Classification, Marking, and Policing Model

Each of these access-edge design options will be discussed in turn.

Conditional Trust Model

As previously mentioned, MQC-based platforms trust at Layer 2 and Layer 3 by default and therefore do not require any explicit commands to perform such functions. Therefore, there are no equivalent commands to mls qos trust cos or mls qos trust dscp (nor are any required).

However, there is a need to provide conditional trust functionality for all switch platforms that may be deployed in the role of an access switch. Hence, there is a corresponding command for conditional trust on the Catalyst 4500 (namely, qos trust device).

At the time of this writing, the Catalyst 4500 supports conditional trust for the following devices:

- Cisco IP phone via the cisco-phone keyword option
- Cisco TelePresence systems via the cts keyword option
- Cisco IP video surveillance cameras systems via the ip-camera keyword option
- Cisco Digital Media Players via the media-player keyword option

When extending conditional trust to Cisco IP phones, it is important to remember that these can only re-mark class of service (CoS) bits (on PC-generated traffic). Therefore, the Conditional Trust Model on the Catalyst 4500 requires a dynamic conditional trust policy applied to the port in conjunction with a simple MQC policy that explicitly matches CoS 5 (for voice) and CoS 3 (for signaling) and marks the DSCP values of these packets to EF and CS3, respectively (essentially performing a CoS-to-DSCP mapping). Example 15-5 shows this conditional trust model for the Catalyst 4500.
Example 15-5  Configuring (CoS-Based) Conditional Trust to a Cisco IP Phone on a Catalyst 4500

! This section defines the class maps to match Voice and Signaling
C4500(config-cmap)# class-map match-all VOICE
C4500(config-cmap)# match cos 5
C4500(config-cmap)# class-map match-all SIGNALING
C4500(config-cmap)# match cos 3

! This section defines the CoS-to-DSCP re-marking policy map
C4500(config-cmap)# policy-map CISCO-IPPHONE
C4500(config-pmap)# class VOICE
C4500(config-pmap-c)# set dscp ef
! Maps CoS 5 to DSCP EF
C4500(config-pmap-c)# class SIGNALING
C4500(config-pmap-c)# set dscp cs3
! Maps CoS 3 to DSCP CS3
C4500(config-pmap-c)# class class-default
C4500(config-pmap-c)# set dscp default
! All other traffic is set to DSCP DF

! This section applies conditional trust and policy map to the int(s)
C4500(config)# interface GigabitEthernet 3/1
C4500(config-if)# switchport access vlan 10
C4500(config-if)# switchport voice vlan 110
C4500(config-if)# spanning-tree portfast
C4500(config-if)# qos trust device cisco-phone
! Applies conditional-trust to the switch port
C4500(config-if)# service-policy input CISCO-IPPHONE
! Attaches the CoS-to-DSCP mapping policy map

You can verify the configuration in Example 15-5 with the following commands:

- show qos interface
- show class-map
- show policy-map
- show policy-map interface
Beginning with Cisco IOS Release IOS XE 3.3.0SG and IOS 15.1(1)SG, you can configure a class map with metadata filters. A QoS policy that includes such classes is termed a metadata-based QoS policy. It allows you to classify flows based on user-friendly metadata attributes rather than on access control list (ACL)-based classification criteria (such as source/destination addresses/ports, and so on).

The following restrictions apply to using a metadata-based QoS policy on a Catalyst 4500 series switch:

- They can only be attached to target in input direction.
- They can only be attached to physical ports and EtherChannel port channel interfaces; they cannot be attached to VLANs, port VLANs, and switch virtual interfaces (SVIs).
- A policy can have multiple metadata-based classifiers.
- A class map can have one or more metadata filters with `match-any` or `match-all` semantics.
- Policy actions corresponding to metadata class are applied on aggregate traffic; however, if the metadata filter is configured along with Flexible NetFlow record filter, the policy action (like policer) applies on individual flows.

**Note** Flow-based QoS policies and Flexible NetFlow (FNF) are discussed further in a following section.

Example 15-6 illustrates a metadata-based QoS policy with two classes using metadata filters.

**Example 15-6 Medianet Metadata Classification Policy Example on a Catalyst 4500**

```
! This section configures the medianet metadata class maps
C4500(config-cmap)# class-map match-all REALTIME-INTERACTIVE
C4500(config-cmap)#  match application telepresence-media
! Identifies TelePresence media flows via metadata
C4500(config-cmap)# class-map match-any MULTIMEDIA-CONFERENCING
C4500(config-cmap)#  match application webex-video
! Identifies WebEx video flows via metadata
C4500(config-cmap)#  match application webex-voice
! Identifies WebEx voice flows via metadata
```

You can verify the configuration in Example 15-6 with the following commands:
Classification and Marking Models

In many scenarios, trust models may not be available or sufficient to distinctly classify all types of traffic required by the end-to-end QoS strategic model. Therefore, explicit classification and marking policies may be needed at the access edge.

Example 15-7 shows a configuration example based on Figure 11-5 (An eight-class QoS model).

Note As previously discussed, not all application classes may be present at the access edge on ingress. For example, streaming video would likely not be present at the access edge on ingress (as these flows are not sourced from campus endpoints, but are likely destined to them), nor would network control flows be sourced from campus endpoints. Therefore, these classes would not need to be included in the access-edge classification and marking policy map.

Note Referenced access lists are omitted from the policy examples for brevity.

Example 15-7 Classification and Marking Policy Example on a Catalyst 4500

```
! This section configures the class maps
C4500(config-cmap)# class-map match-all VOICE
C4500(config-cmap)# match dscp ef
! Voice is matched on DSCP EF
C4500(config-cmap)# class-map match-all INTERACTIVE-VIDEO
C4500(config-cmap)# match access-group name INTERACTIVE-VIDEO
! Associates INTERACTIVE-VIDEO access-list with class map
C4500(config-cmap)# class-map match-all SIGNALING
C4500(config-cmap)# match cs3
! Signaling is matched on DSCP CS3
C4500(config-cmap)# class-map match-all TRANSACTIONAL-DATA
C4500(config-cmap)# match access-group name TRANSACTIONAL-DATA
! Associates TRANSACTIONAL-DATA access-list with class map
C4500(config-cmap)# class-map match-all SCAVENGER
C4500(config-cmap)# match access-group name SCAVENGER
! Associates SCAVENGER access-list with class map
```
! This section configures the Per-Port ingress marking policy map
C4500(config-cmap)# policy-map PER-PORT-MARKING
C4500(config-pmap)# class VOICE
C4500(config-pmap-c)# set dscp ef
! VoIP is marked EF
C4500(config-pmap-c)# class INTERACTIVE-VIDEO
C4500(config-pmap-c)# set dscp af41
! Interactive-Video is marked AF41
C4500(config-pmap-c)# class SIGNALING
C4500(config-pmap-c)# set dscp cs3
! Signaling is marked CS3
C4500(config-pmap-c)# class TRANSACTIONAL-DATA
C4500(config-pmap-c)# set dscp af21
! Transactional Data is marked AF21
C4500(config-pmap-c)# class SCAVENGER
C4500(config-pmap-c)# set dscp cs1
! Scavenger traffic is marked CS1
C4500(config-pmap-c)# class class-default
C4500(config-pmap-c)# set dscp default
! All other traffic is marked DF

! This section attaches the service-policy to the interface(s)
C4500(config)# interface range GigabitEthernet 2/1-48
C4500(config-if-range)# switchport access vlan 10
C4500(config-if-range)# switchport voice vlan 110
C4500(config-if-range)# spanning-tree portfast
C4500(config-if-range)# qos trust device cisco-phone
! The interface is set to conditionally trust Cisco IP Phones
C4500(config-if-range)# service-policy input PER-PORT-MARKING
! Attaches the Per-Port Marking policy to the interface(s)

You can verify the configuration in Example 15-7 with the following commands:

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

Classification, Marking, and Policing Model
In addition to classification and marking, policing might also be required at the access edge. The Catalyst 4500 can perform single-rate (two-color) policing and three-color...
policing—via either the RFC 2697 single-rate three-color marker (srTCM) or the RFC 2698 two-rate three-color marker (trTCM). Example 15-8 shows a per-port single-rate policing example for the Catalyst 4500 (based on Figure 13-8), and Example 15-9 shows policy amendments to support a RFC 2698 two-rate three-color marker.

**Example 15-8** *(Single-Rate Two-Color) Per-Port Policing Configuration Example on a Catalyst 4500*

```
! This section configures the single-rate per-port policing policy map
C4500(config)# policy-map PER-PORT-POLICING
C4500(config-pmap)# class VVLAN-VOIP
C4500(config-pmap-c)# set dscp ef
C4500(config-pmap-c)# police 128k bc 8000
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action drop
! VoIP is marked EF and policed to drop at 128 kbps
C4500(config-pmap)# class VVLAN-SIGNALING
C4500(config-pmap-c)# set dscp cs3
C4500(config-pmap-c)# police 32k bc 8000
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action drop
! (VVLAN) Signaling is marked CS3 and policed to drop at 32 Kbps
C4500(config-pmap)# class MULTIMEDIA-CONFERENCING
C4500(config-pmap-c)# set dscp af41
C4500(config-pmap-c)# police 5m bc 8000
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action drop
! Multimedia-conferencing is marked AF41 and policed to drop at 5 Mbps
C4500(config-pmap)# class SIGNALING
C4500(config-pmap-c)# set dscp cs3
C4500(config-pmap-c)# police 32k bc 8000
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action drop
! (DVLAN) Signaling is marked CS3 and policed to drop at 32 Kbps
C4500(config-pmap)# class TRANSACTIONAL-DATA
C4500(config-pmap-c)# set dscp af21
C4500(config-pmap-c)# police 10m bc 8000
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action set-dscp-transmit af22
! Trans-data is marked AF21 and policed to re-mark (to AF22) at 10 Mbps
C4500(config-pmap)# class BULK-DATA
C4500(config-pmap-c)# set dscp af11
C4500(config-pmap-c)# police 10m bc 8000
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action set-dscp-transmit af12
```
Bulk-data is marked AF11 and policed to re-mark (to AF12) at 10 Mbps

```
C4500(config-pmap)# class SCAVENGER
C4500(config-pmap-c)# set dscp cs1
C4500(config-pmap-c)# police 10m bc 8000
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action drop

! Scavenger traffic is marked CS1 and policed to drop at 10 Mbps
```

```
C4500(config-pmap)# class class-default
C4500(config-pmap-c)# set dscp default
C4500(config-pmap-c)# police 10m bc 8000
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action set-dscp-transmit cs1

! The implicit default class marks all other traffic to DF
! and polices all other traffic to re-mark (to CS1) at 10 Mbps
```

```
! This section attaches the service-policy to the interface(s)
C4500(config)# interface range GigabitEthernet 2/1-48
C4500(config-if-range)# switchport access vlan 10
C4500(config-if-range)# switchport voice vlan 110
C4500(config-if-range)# spanning-tree portfast
C4500(config-if-range)# qos trust device cisco-phone
! The interface is set to conditionally trust Cisco IP phones
C4500(config-if-range)# service-policy input PER-PORT-POLICING
! Attaches the Per-Port Policing policy to the interface(s)
```

Note  The Catalyst 4500 IOS Software allows for policing rates to be entered using the postfixes k (for kilobits), m (for megabits), and g (for gigabits), as shown in Example 15-8. In addition, decimal points are allowed in conjunction with these postfixes. For example, a rate of 10.5 Mbps could be entered with the policy map command `police 10.5m`. These policing rates are converted to their full bits-per-second values within the configuration, but it makes the entering of these rate more user friendly and less error prone (as could easily be the case when having to enter up to 10 zeros to define the policing rate).

You can verify the configuration in Example 15-8 with the following commands:

- `show qos interface`
- `show class-map`
- `show policy-map`
- `show policy-map interface`
To avoid excessive repetition, Example 15-9 amends and expands the policer from a single-rate two-color marker to a two-rate three-color marker only on a single class (the Bulk Data class). However, similar amendments can be made on any Assured Forwarding (AF) class of traffic.

**Example 15-9  (Two-Rate Three-Color) Per-Port Policing Configuration Amendment Example on a Catalyst 4500**

```plaintext
! This section configures a dual-rate per-port policing policy map
C4500(config)# policy-map TWO-RATE-POLICER

<snip>

C4500(config-pmap)# class BULK-DATA
C4500(config-pmap-c)# set dscp af11
C4500(config-pmap-c)# police 10m bc 8000 pir 15m
! Bulk-data is policed to 10 Mbps rate and 15 Mbps peak rate
C4500(config-pmap-c-police)# conform-action set-dscp-transmit af11
! Bulk data under 10 Mbps will be marked AF11
C4500(config-pmap-c-police)# exceed-action set-dscp-transmit af12
! Bulk data traffic between 10 Mbps and 15 Mbps will be marked AF12
C4500(config-pmap-c-police)# violate-action set-dscp-transmit af13
! Bulk data traffic over 15Mbps will be marked AF13
```

You can verify the configuration in Example 15-9 with the following commands:

- `show qos interface`
- `show class-map`
- `show policy-map`
- `show policy-map interface`

**Per-VLAN QoS Design**

The Catalyst 4500 supports VLAN-based QoS. However, unlike the Catalyst 3750, the Catalyst 4500 does not support the `mls qos vlan-based` interface command. Furthermore, service policies are attached to VLANs via the VLAN configuration mode (instead of the interface configuration mode), as shown in Example 15-10.

**Example 15-10  Per-VLAN Marking Configuration Example on a Catalyst 4500**

```plaintext
! This section configures the interface(s) for conditional trust,
C4500(config)# interface range GigabitEthernet 2/1-48
C4500(config-if-range)# switchport access vlan 10
```
C4500(config-if-range)# switchport voice vlan 110
C4500(config-if-range)# spanning-tree portfast
C4500(config-if-range)# qos trust device cisco-phone

! The interface is set to conditionally trust Cisco IP phones

! This section attaches a marking policy to the DVLAN
C4500(config)# vlan config 10
C4500(config-vlan-config)# service-policy input DVLAN-MARKING

! This section attaches a marking policy to the VVLAN
C4500(config)# vlan config 110
C4500(config-vlan-config)# service-policy input VVLAN-MARKING

You can verify the configuration in Example 15-10 with the following commands:

- show qos interface
- show class-map
- show policy-map
- show policy-map vlan vlan-number (This command is nearly identical to show policy map interface, except that it references a VLAN directly, rather than a VLAN interface.)

**Note** It is not recommended to deploy policing policies on a per-VLAN basis, as discussed further in the next section.

**Per-Port/Per-VLAN QoS**

Although it is technically possible to apply a (aggregate) policing policy on a per-VLAN basis, it is not advisable to do so. This is because the number of endpoints in a given VLAN can dynamically vary, yet the policing rates are statically fixed at an aggregate level, resulting in unpredictable bandwidth allotments per endpoint.

However, a more flexible and discrete approach for deploying policing policies exists on the Catalyst 4500 platforms—namely, to deploy these on a per-port/per-VLAN basis. The Catalyst 4500 has a very elegant syntax for deploying per-port/per-VLAN policies, as follows: Within a (trunked) switch port's interface configuration, each VLAN carried over that trunked port can have a separate policy applied to it via an interface-vlan configuration mode, as shown in Example 15-11.
Example 15-11  Per-Port/Per-VLAN Policing Configuration Example on a Catalyst 4500

```cisco
! This section attaches the policy to the VLANS on a per-port basis
C4500(config)# interface range GigabitEthernet 2/1-48
C4500(config-if-range)# switchport access vlan 10
C4500(config-if-range)# switchport voice vlan 110
C4500(config-if-range)# spanning-tree portfast
C4500(config-if-range)# qos trust device cisco-phone
! The interface is set to conditionally trust Cisco IP phones
C4500(config-if-range)# vlan 10
C4500(config-if-vlan-range)# service-policy input DVLAN-POLICERS
! Attaches the per-port/per-VLAN DVLAN policing policy to the
! DVLAN of the trunked switch port(s)
C4500(config-if-range)# vlan 110
C4500(config-if-vlan-range)# service-policy input VVLAN-POLICERS
! Attaches the per-port/per-VLAN VVLAN policing policy to the
! VVLAN of the trunked switch port(s)
```

You can verify the configuration in Example 15-11 with the following commands:

- show qos interface
- show class-map
- show policy-map
- show policy-map interface
- show policy-map interface interface x/y vlan vlan-number

EtherChannel QoS Design

The following rules apply when deploying QoS service policies on Catalyst 4500 EtherChannels:

- Classification, marking, and policing policies (whether ingress or egress) are applied to the logical port channel interfaces.
- Queuing policies are applied to the physical port-member interfaces.

For EtherChannel interfaces configured on Catalyst 4500 switches, the ingress QoS policies (including classification, marking, and policing policies) are applied via MQC service-policy statements (in the ingress direction using the input keyword) configured on the logical port channel interface. Trust statements are not required because this MQC-based platform trusts by default.
In addition, the Catalyst 4500 supports **egress QoS policies** (including marking/policing policies) to be similarly applied via MQC **service-policy** statements (in the egress direction using the **output** keyword) on the **logical port channel interface**.

**Egress queuing** policies, however, are applied via MQC **service-policy** statements (in the egress direction using the **output** keyword) on the **physical port-member interfaces**, as shown in Example 15-12.

**Example 15-12  EtherChannel QoS Design on a Catalyst 4500**

```
! This section configures the logical port channel interface
C4500(config)# interface Port-channel1
C4500(config-if)# description ETHERCHANNEL-LOGICAL-INTERFACE
C4500(config-if)# switchport mode trunk
C4500(config-if)# switchport trunk encapsulation dot1q
C4500(config-if)# switchport trunk allowed vlan 10,110
C4500(config-if)# service-policy input MARKING

! This section configures 1P3Q1T+DBL queuing on physical port-member interfaces
C4500(config)# interface range TenGigabitEthernet1/1-2
C4500(config-if-range)# description PORT-CHANNEL1-PORT-MEMBER
C4500(config-if-range)# switchport mode trunk
C4500(config-if-range)# switchport trunk encapsulation dot1q
C4500(config-if-range)# switchport trunk allowed vlan 10,110
C4500(config-if-range)# channel-group 1 mode auto
C4500(config-if-range)# service-policy output 1P7Q1T-QUEUING
! Applies 1P7Q1T+DBL-QUEUING queuing policy to physical port member
```

You can verify the configuration in Example 15-12 with the following commands:

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

**Note** As previously stated, the queueing policies will only attach to EtherChannel port-member physical interfaces if the priority queue is not explicitly policed. If policing the priority queue is desired, a separate policy map needs to be constructed to do so and attached to the logical EtherChannel interface in the **egress** direction.
Flow-Based QoS

Flow-based QoS enables microflow policing and marking capability to dynamically learn traffic flows, providing the capability to police every unique flow to an individual rate. Flow-based QoS is available on a Catalyst 4500 series switch with the built-in NetFlow hardware support. It can be applied to ingress traffic on both switched and routed interfaces with flow masks defined using Flexible NetFlow (FNF). Flow-based QoS is typically used in environments where per-user, granular rate limiting is required. Flow-based QoS is also referred to as user-based rate limiting (UBRL).

A flow is defined as a stream of packets having the same properties as those defined by the key fields in the FNF flow record. A new flow is created when the value of data in packet’s key fields is unique with respect to the flows that already exist.

A flow-based QoS policy is possesses one or more class maps matching on a FNF flow record. Such a class map must be configured as match-all to match all the match criteria specified in the class map. When a flow-based QoS policy is attached to a QoS target, ingress traffic on the target is first classified based on the classification rules specified in the class map. If the classifier has an FNF flow record, the key fields specified in the FNF flow record are applied on the classified traffic to create flows provided the flow does not already exist. The corresponding policy actions (policing and marking) are then applied to these individual flows. Flow-based policers (termed microflow policers) rate limit each unique flow. Flows are dynamically created and inactive flows are periodically aged out.

Flow-based QoS policy can be applied on a per-port basis, per-port/per-VLAN basis, or on an EtherChannel port channel interface (but only in the ingress direction). Therefore, flow-based QoS may be deployed at either the access layer or distribution layer (wherever UBRL may be of value).

Note that flow-based policies will apply to all flows matched within a given class. For example, if a flow-based policer is applied to the default class and attached to port or VLAN, all flows originating from that port or VLAN (respectively) will be subject to the policer. If this is not to be the intent, additional classification is recommended and the flow-based policer should be more selectively applied.

Example 15-13 shows how to configure a flow-based QoS policy that uses microflow policing in the context of user-based rate limiting. Any and all flows sourced from the subnet 192.168.10.* are microflow policed to 1 Mbps.

Example 15-13  Configuring Flow-Based QoS (UBRL) on Catalyst 4500

```bash
! This section defines an ACL to match traffic from subnet
C4500(config)# ip access-list extended USERGROUP-1
C4500(config-ext-nacl)# permit ip 192.168.10.0 0.0.0.255 any
! Traffic sourced from the 192.168.10.x subnet is matched
```
This section defines a flow record with source address as key.

```
C4500(config)# flow record FLOW-RECORD-1
C4500(config-flow-record)# match ipv4 source address
```

Source address is defined as the key tuple.

This section defines the class map to match on USERGROUP-1 ACL and specify FLOW-RECORD-1 definition for flow creation.

```
C4500(config)# class-map match-all USER-GROUP-1
C4500(config-cmap)# match access-group name USERGROUP-1
C4500(config-cmap)# match flow record FLOW-RECORD-1
```

A "match-all" class map binds the ACL and flow-record to identify unique flows.

This section defines the microflow policer policy map.

```
C4500(config)# policy-map 1MBS-MICROFLOW-POLICER
C4500(config-pmap)# class USER-GROUP-1
C4500(config-pmap-c)# police cir 1m
C4500(config-pmap-c-police)# conform-action transmit
C4500(config-pmap-c-police)# exceed-action drop
```

Specifies each discrete microflow is to be limited to 1Mbs.

This section applies the microflow policer to the interface.

```
C4500(config)# interface gigabitEthernet3/1
C4500(config-if)# service-policy input 1MBS-MICROFLOW-POLICER
```

You can verify the configuration in Example 15-13 with the following commands:

- show flow record (demonstrated in Example 15-14)
- show class-map
- show policy-map
- show policy-map interface

**Example 15-14  Verifying Flow-Based QoS Policies on a Catalyst 4500: show flow record**

```
C4500# show flow record
flow record FLOW-RECORD-1:
  Description: User defined
  No. of users: 1
```
AutoQoS SRND4

AutoQoS SRND4 is supported on the Cisco Catalyst 4500 beginning with Cisco IOS Release IOS XE 3.3.0SG and IOS 15.1(1)SG and is detailed in Appendix A, “AutoQoS for Medianet.”

Control Plane Policing

Control plane policing (CPP) is supported on the Catalyst 4500 and is detailed in Appendix B, “Control Plane Policing.”

Summary

This design chapter primarily discussed the best-practice QoS design recommendations for the Cisco Catalyst 4500 (Supervisor 6-E/7-E) series switch in the role of a campus distribution layer switch. (which, incidentally are equivalent to the QoS designs required were it serving in the role of a campus core switch).

Because the Catalyst 4500 is an MQC-based QoS platform, QoS is enabled by default, as is DSCP trust, on all ports. Therefore, there is effectively only a single step to configuring QoS on a Catalyst 4500 performing the role of a distribution switch: to configure an egress queuing policy.

To this end, 4-class, 8-class, and 12-class queuing policies were detailed, along with corresponding configurations and verification examples, leveraging the Catalyst 4500’s flexible 1P7Q1T+DBL hardware queuing capabilities.

Additional platform-specific design options and considerations were discussed, including how the Catalyst 4500 could be deployed as an access-edge switch, and how to configure per-VLAN QoS, per-port/per-VLAN QoS, and EtherChannel QoS designs.

AutoQoS SRND4 is supported on the Catalyst 4500 and is covered in Appendix A; similarly, CPP is also supported and is covered in Appendix B.

Further Reading


### A

<table>
<thead>
<tr>
<th>Access/Aggregation Layer Nexus 5500/2000 QoS Design (Tifosi Software Inc. Case Study)</th>
<th>659-666</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access-Edge Design</td>
<td></td>
</tr>
<tr>
<td>Campus Core (Cisco Catalyst 6500) QoS Design</td>
<td></td>
</tr>
<tr>
<td>Classification, Marking, and Policing Models, 335-340</td>
<td></td>
</tr>
<tr>
<td>Classification and Marking Models, 332-335</td>
<td></td>
</tr>
<tr>
<td>Conditional Trust Models, 330-332</td>
<td></td>
</tr>
<tr>
<td>Overview, 330</td>
<td></td>
</tr>
<tr>
<td>Cisco Catalyst 4500</td>
<td></td>
</tr>
<tr>
<td>Classification, Marking, and Policing Model, 295-297</td>
<td></td>
</tr>
<tr>
<td>Classification and Marking Models, 293-294</td>
<td></td>
</tr>
<tr>
<td>Conditional Trust Model, 290-291</td>
<td></td>
</tr>
<tr>
<td>Medianet Metadata Classification Model, 292-293</td>
<td></td>
</tr>
<tr>
<td>Overview, 290</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tifosi Software, Inc. (Case Study)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco IP Phones and PCs (Conditional Trust and Classification and Marking), 482-485</td>
<td></td>
</tr>
<tr>
<td>Cisco TelePresence Endpoints (Conditional Trust), 482</td>
<td></td>
</tr>
<tr>
<td>Mobile Wireless Clients (Dynamic Policy with Classification and Marking), 489-490</td>
<td></td>
</tr>
<tr>
<td>Wired Access Endpoints (DSCP Trust), 481-482</td>
<td></td>
</tr>
<tr>
<td>Wired Printer Endpoints (No Trust), 481</td>
<td></td>
</tr>
<tr>
<td>Wired Queuing, 485-488</td>
<td></td>
</tr>
<tr>
<td>Wireless Queuing, 491-492</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Access Layer Uplink Design (Tifosi Software Inc. Case Study), 359-360</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ACs (Access Categories), 383-385</td>
<td></td>
</tr>
<tr>
<td>Address-Based Classifications, 19-20</td>
<td></td>
</tr>
<tr>
<td>Admission Control, 14, 100-101</td>
<td></td>
</tr>
<tr>
<td>Advanced RSVP Model with Application ID, 729-733</td>
<td></td>
</tr>
<tr>
<td>AF (Assured Forwarding), 685</td>
<td></td>
</tr>
<tr>
<td>AF Queue Recommendations, 195</td>
<td></td>
</tr>
<tr>
<td>AIFS (Arbitration Interframe Spacing Number), 385-386</td>
<td></td>
</tr>
</tbody>
</table>
AP (access point), 4
application-based classifications, 19-20, 739-743, 745-747
application class expansion QoS strategies, 204-205
application-group-based classification model, 743-744, 748
application policing server models, 578-580
application trends
control plane traffic, 180-182
data applications, 177-180
multimedia applications, 175-177
overview, 169-170
video applications, 171-175
voice, 170-171
architecture. See also specific architectures
data center access/aggregation (Nexus 5500/2000) QoS design overview, 562-564 QoS groups and system classes, 567-569 QoS policies supported by, 562-564 VOQ (virtual output queuing), 564-567

data center QoS design considerations and recommendations
big data (HPC/HTC/Grid) architectures, 501-502
high-performance trading data center architectures, 500-501
massively scalable data center architectures, 506
overview, 500
secure multitenant data center architectures, 505
virtualized multiservice data center architectures, 503-505
data center virtual access (Nexus 1000V) QoS design, 537-539 Medianet, 119-120 QoS, 14-16
service provider core (Cisco CRS) QoS design, 846-849
service provider edge (Cisco ASR 9000) QoS design, 810-814 ASR (Aggregation Services Routers), 190
assured forwarding (AF), 685
asymmetrical CoS/DSCP marking, 526
ATM (Asynchronous Transfer Mode), 3, 38-39
ATM traffic shaping, 78
attribute-based classification model, 744, 748-752
auto qos classify, 28
auto qos trust, 28
auto qos video, 28
auto qos voip, 28
Auto Smartports, 121, 243
autodiscovery (data collection), 28
AutoQoS
marking, 54
Medianet
Cisco Catalyst 4500 series switches, 971-982
classify and police models, 958-963
overview, 953-955
1P3Q3T egress queuing models, 969-971
1P1Q3T ingress queuing models, 968-969
trust models, 955-956
video models, 956-958
VoIP models, 963-968
overview, 25-28, 121-122
SRND4
  branch router (Cisco ISR G2)
  QoS design, 757
  campus access (Cisco Catalyst 3750) QoS design, 274
  Cisco Catalyst 4500, 303
  WAN aggregator (Cisco ASR 1000) QoS design, 708, 733
VoIP, 27, 242
AVC (application visibility control)
ASR 1000 routers, 137
building blocks, 140-159
Cisco wireless LAN routers, 137
FNF (Flexible NetFlow)
  configuration, 149-152
  key fields, 148-149
  non-key fields, 148-149
  overview, 147-148
  performance considerations, 159-160
how it works, 138-140
Internet edge, 137
ISR G2 routers, 137
management and reporting
  Insight Reporter, 153
  overview, 152-153
NBAR2
  MQC classification, 144-147
  overview, 140-142
  performance considerations, 159-160
  protocol discovery, 142-144
overview, 136-137
performance considerations, 159-160
QoS controls
  Internet edge, deploying AVC
  QoS controls at, 156-158
  overview, 154
  WAN edge, deploying AVC
  QoS controls at, 154-156
  use cases, 136-137
  WAN and branch QoS design considerations and recommendations, 687
  WAN edge, 137
  wireless LAN controller (Cisco 5500) QoS design, 417-424

B

bandwidth
  allocation, 14
  changes in, 2
bandwidth reservation tools
  admission control tools, 100-101
  overview, 99-100
  recommendations and guidelines, 108
RSVP
  deployment models, 103-106
  LLQ and, 106-107
  overview, 101-102
  proxy, 102-103
basic RSVP model, 726-729
behavioral model for QoS, 15
best effort data, 179
best practice design principles
  classification and marking best practices, 191-192
  hardware versus software QoS best practices, 190
overview, 189-190
policing and markdown best practices, 192
queuing and dropping best practices
  AF queue recommendations, 195
  DF queue recommendations, 195
  EF queue recommendations: the 33% LLQ rule, 193-195
  overview, 192-193
  scavenger class queue recommendations, 195-196
  WRED recommendations, 197
big data (HPC/HTC/Grid) architectures, 501-502
bottom-up applications, 168
branch LAN edge, 693
branch router (Cisco ISR G2) QoS design
egress QoS models
  eight-class model, 754
  four-class model, 754
  overview, 753
  twelve-class model, 754-756
ingress QoS models
  Medianet classification models, 738-744
  NBAR2 classification models, 744-753
  overview, 738
overview, 753, 757
platform-specific QoS design options
  AutoQoS SRND4, 757
  control plane policing, 757
  overview, 757
  RSVP, 757
branch routers, 677-678
branch WAN edge, 693
c broadcast streams, 165
broadcast video, 34, 173-174
Bronze QoS profile for wireless
  LAN controller (Cisco 5500) QoS design, 400-408
buffer size, modifying ingress, 580-582
bulk data (high-throughput data), 178-179
business and application QoS requirements
  application trends
    control plane traffic, 180-182
    data applications, 177-180
    multimedia applications, 175-177
      overview, 169-170
    video applications, 171-175
    voice, 170-171
bottom-up applications, 168
BYOD (bring your own device), 167-168
global trends in networking, 164
high-definition media, 169
media content, increase in, 166-167
multimedia applications, convergence of media subcomponents within, 168-169
QoS standards evolution
  overview, 183
  RFC 2597 (clarification), 183-184
  RFC 4594 (update draft), 185-187
  RFC 5865 (proposed standard), 184-185
RFC 4594-based application class QoS recommendations, 182
social networking, appearance and effect on business networks of, 167
top-down deployments, 168
video applications, evolution of, 164-166
business catalysts for QoS reengi-neering (Tifosi Software Inc. case study), 216-217
BYOD (bring your own device), 167-168

C

C-Vision, 167
CAC (call admission control)
overview, 62, 99-100
wireless LAN controller (Cisco 5500) QoS design
configuration, 414-415
overview, 413
campus access (Cisco Catalyst 3750) QoS design
Cisco Catalyst 3750 QoS architecture, 248-249
classification, marking, and policing models, 256-259
classification and marking models, 254-256
enabling QoS globally, 250
ingress QoS models, configuring, 250-259
overview, 247-248
platform-specific QoS design options
AutoQoS SRND4, 274
EtherChannel QoS design, 273 overview, 271
per-port/per-VLAN QoS design, 272-273
per-VLAN QoS design, 271-272
queuing models
egress queuing model, 265-271
ingress queuing model, 261-265
overview, 260-261
steps for, 249-271
trust models
conditional trust models, 253-254
overview, 251
trust CoS model, 251-252
trust DSCP model, 252
untrusted model, 251
campus AutoQoS, 241-243
campus CE ingress/internal QoS (ASR 1000), 788
campus core (Cisco Catalyst 6500) QoS design
access-edge design options
classification, marking, and policing models, 335-340
classification and marking models, 332-335
conditional trust models, 330-332
overview, 330
architecture, 306-308
overview, 305-306
platform-specific QoS design options
access-edge design options, 330-340
CPP (control plane policing), 344
EtherChannel QoS design, 343-344
microflow policing, 341-342
overview, 329-330
per-VLAN QoS design, 342-343
queuing models
eight-class (8Q4T ingress and 1P7Q4T egress) queuing models, 314-318
four-class (4Q4T ingress and 1P3Q4T egress) queuing models, 311-314
overview, 308-311
2P6Q4T ingress and egress queuing models, 328-329
twelve-class (8Q4T ingress and 1P7Q4T egress) queuing models, 318-328
steps for, 308
campus distribution (Cisco Catalyst 4500) QoS design
Cisco Catalyst 4500 QoS architecture, 276-277
configuring QoS on Cisco Catalyst 4500, 277
overview, 275
platform-specific QoS design options
access-edge design options, 290-297
AutoQoS SRND4, 303
CPP (control plane policing), 303
EtherChannel QoS design, 299-300
flow-based QoS design, 301-303
overview, 289
per-port/per-VLAN QoS design, 298-299
per-VLAN QoS design, 297-298
queuing models
eight-class egress queuing model, 281-284
four-class egress queuing model, 278-281
overview, 277-278
twelve-class egress queuing model, 284-289
campus port QoS roles
overview, 239
switch ports connecting to conditionally trusted endpoints, 240
switch ports connecting to network infrastructure, 241
switch ports connecting to trusted endpoints, 240
switch ports connecting to untrusted endpoints, 240
campus QoS design (Tifosi Software Inc. case study)
access layer uplink design, 359-360
access QoS design, 350-360
Catalyst 3750, 350-360
Catalyst 4550, 360-364
Catalyst 6550, 364-370
Cisco IP phones or PCs (conditional trust and classification and marking), access-edge design for, 352-355
Cisco TelePresence endpoints (conditional trust), access-edge design for, 352
core layer (40GE) core-link design, 368-370
core layer (10GE) downlink design, 364-368
core QoS design, 364-370
distribution layer distribution-link/core-uplink ports, 362-364
distribution layer downlink ports, 360-362
distribution QoS design, 360-364
eight-class 1P3Q3T egress queuing design, 357-359
eight-class 1P1Q3T ingress queuing design, 355-357
overview, 347-350
printer endpoints, access-edge design for, 351
wireless access endpoints (DSCP Trust), access-edge design for, 351

campus QoS design considerations and recommendations
AutoQoS, 241-243
CoPP (control plane policing), 243-244
default QoS, 226
DSCP transparency, 231
EtherChannel QoS, 234-235
internal DSCP, 226-227
MLS versus MQC, 225-226
overview, 223-225
port-based QoS versus VLAN-based QoS versus per-port/per-VLAN QoS, 232-233
port QoS roles
overview, 239
switch ports connecting to conditionally trusted endpoints, 240
switch ports connecting to network infrastructure, 241
switch ports connecting to trusted endpoints, 240
switch ports connecting to untrusted endpoints, 240
QoS models
egress QoS models, 238-239
ingress QoS models, 235-237
overview, 235
trust boundaries, 230-231
trust states and operations, 227-230
CAPWAP (Control and Wireless Access Points), 40, 389
CBWFQ (class-based weighted fair queuing), 87-89
scavenger CBWFQs, 691
WAN and branch QoS design considerations and recommendations, 683
CE LAN edge, 788
CE routers (Tifosi Software Inc. case study)
internal QoS (Cisco ASR 1000), 863
LAN-edge QoS policies, 863
VPN-edge QoS policies, 863-866
CE VPN edge, 788
circuit-switched networks, 3
Cisco ASR 9000 QoS design
architecture, 810-814
MPLS DiffServ tunneling models
overview, 814-815
pipe mode MPLS DiffServ tunneling, 826-834
short pipe mode MPLS DiffServ tunneling, 834-842
uniform mode MPLS DiffServ tunneling, 815-826
overview, 809
steps for, 814
Cisco ASR 1000 routers. See also WAN aggregator (Cisco ASR 1000) QoS design, 708, 733
AVC (application visibility control), 137
internal QoS
  overview, 701
SIP-based PLIM, 707-708
SIP-10s oversubscription scenarios, 703
SPA-based matrix of ingress classification by SIP or SPA level, 705-706
SPA-based PLIM,
Cisco Catalyst 3750 (Tifosi Software Inc. case study), 350-360. See also campus access (Cisco Catalyst 3750) QoS design
Cisco Catalyst 3850. See also converged access (Cisco Catalyst 3850 and Cisco 5760 Wireless LAN controller) QoS design
CPP/CoPP (control plane policing), 987-990
trust policy, 443-444-446
Cisco Catalyst 4500
  access-edge design options
    classification, marking, and policing model, 295-297
    classification and marking models, 293-294
    conditional trust model, 290-291
    Medianet metadata classification model, 292-293
    overview, 290
configuring QoS on Cisco Catalyst 4500, 277
CPP/CoPP (control plane policing), 989-996
overview, 275
platform-specific QoS design options
  access-edge design options, 290-297
AutoQoS SRND4, 303
CPP (control plane policing), 303
EtherChannel QoS design, 299-300
flow-based QoS design, 301-303
overview, 289
per-port/per-VLAN QoS design, 298-299
per-VLAN QoS design, 297-298
QoS architecture, 276-277
queuing models
  eight-class egress queuing model, 281-284
  four-class egress queuing model, 278-281
  overview, 277-278
  twelve-class egress queuing model, 284-289
Cisco Catalyst 4550 (Tifosi Software Inc. case study), 360-364
Cisco Catalyst 6500, 996-998. See also campus core (Cisco Catalyst 6500) QoS design
Cisco Catalyst 6550 (Tifosi Software Inc. case study), 364-370
Cisco Catalyst 3650-E/X, 248-249
Cisco Catalyst 2960-G, 248-249
Cisco Catalyst 2960-G/S, 248-249
Cisco Catalyst 2960-S, 248-249
Cisco Catalyst 4500 series switches, 971-982
Cisco CRS QoS design
  architecture, 846-849
design steps, 849
overview, 845-846
SP core CoS QoS models
  eight-CoS SP core model, 857-860
  four-CoS SP model, 850-854
  overview, 849-850
  six-CoS SP core model, 854-857
Cisco 5500 wireless LAN controllers
  AVC (application visibility control), 417-424
  Bronze QoS profile, 400-408
  CAC (call admission control)
    configuring, 414-415
    overview, 413
  downstream traffic, 425-429
  EDCA, optimizing, 411-412
  eight-class model design, 430-431
  enforcement points, 398
  four-class model design, 425-430
  Gold QoS profile, 400-408
  guest QoS profile, building, 408-410
  Media Session (SIP) snooping, 416-417
  overview, 397
  Platinum QoS profile, 400-408
  Silver QoS profile, 400-408
  strategy, developing, 424-431
  trust boundaries, 399-400
  twelve-class model design, 431
  upstream traffic, 429-430
  VoIP applications, 410-413
  WLAN QoS profiles, 400-408
  WMM policy
    enabling, 413-414
    overview, 405-408
Cisco IP phones or PCs (conditional trust and classification and marking), access-edge design for, 352-355
Cisco ISE (Identity Services Engine), 495
Cisco ISR G2 QoS design, 738-739, 744-745
Cisco Nexus 7000
  F2/F2e-Series I/O modules
    additional design options, 638-648
    architecture, 623-625
    default network QoS policy design, 625-629
    FEX (Fabric Extender) QoS design, 638
    overview, 630
    QoS design steps, 625
    queuing models, 630-637
  fabric modules, 600
  M2-Series I/O modules
    additional design options, 638-648
    architecture, 604-607
    OTV (Overlay Transport Virtualization) edge device QoS design, 621-623
    overview, 607
    QoS design steps, 607
    queuing models, 607-621
  overview, 600-604
  QoS policies supported by, 601-602
  supervisor modules, 600
  trust default behavior, 602-603
Cisco Nexus 2000 fabric extender QoS, 593-596
Cisco Nexus OS QoS framework, 519-520
Cisco Nexus 1000V (data center virtual access) QoS design
architecture, 537-539
configuration notes, 539-540
egress QoS models
  eight-class queuing model, 556-558
  four-class queuing model, 551-556
  overview, 549-551
ingress QoS models
  classification and marking, 544-547
  overview, 541
  server policing model, 547-549
  trusted models, 541-544
  overview, 535-537
statistics, monitoring QoS, 540
trust models
  trusted server model, 541
  untrusted server model, 541-544
VEM (virtual ethernet module), 537-539
VSM (virtual supervisor module), 537-539
Cisco TelePresence, 166, 169, 352
Cisco to RFC 4594 markings, mapping, 42
Cisco Unified Communications Manager (CUCM), 103
Cisco Unified Wireless Networking (CUWN), 435-436
Cisco Visual Networking Index: Forecast and Methodology Report, 164
Cisco wireless LAN routers, 137
class-map command, 17
class maps
  addressing information, 19-20, 46
application-based classifications, 19-20
feature sequence, effects of, 52
logical or physical interface, 46
marking-based classifications, 19-20
MQC (modular QoS command-line) framework, 19-20
overview, 50-52
packet attributes, characteristics, or field values, 45
packet discard eligibility, 51
packet header markings, 45
ports, 46
protocols, 45-46
table map feature, mapping markings with, 52-53
ToS values, 51
tunnel ToS values, 51
classification
  defined, 32
  QoS, 14-15
classification, marking, and policing models
campus access (Cisco Catalyst 3750) QoS design, 256-259
campus core (Cisco Catalyst 6500) QoS design, 335-340
Cisco Catalyst 4500, 295-297
converged access (Cisco Catalyst 3850 and Cisco 5760 Wireless LAN controller) QoS design, 448-454
classification and marking
  best practices, 191-192
campus access (Cisco Catalyst 3750) QoS design, 254-256
campus core (Cisco Catalyst 6500) QoS design, 332-335
Cisco Catalyst 4500, 293-294
converged access (Cisco Catalyst 3850 and Cisco 5760 Wireless LAN controller) QoS design, 446-448

data center access/aggregation (Nexus 5500/2000) QoS design, 572-578

defined, 32

defining QoS markings

Cisco to RFC 4594 markings, mapping, 42
L2 to L3 markings, mapping, 41-42
overview, 41
wireless networks, mapping markings for, 43

marking fields in different technologies

ATM, 38-39
CAPWAP, 40
Ethernet 802.11 WiFi, 38
Ethernet 802.1Q/p, 37
field values and interpretation, 35-37
FR, 38-39
IPv4, 39
IPv6, 39
L2 tunnels, 40
L3 tunnels, 40
MPLS, 41
overview, 35

recommendations and guidelines, 55

security
network attacks, 34
trust boundaries, 33
terminology, 32-33
tools, 7
video traffic, 34

wireless traffic, 35

classification tools
class-based classification (class maps), 45-47
addressing information, 46
logical or physical interface, 46
packet attributes, characteristics, or field values, 45
packet header markings, 45
ports, 46
protocols, 45-46

NBAR (network based application recognition)
metadata classification, 50
overview, 47-48
performance routing, 49-50
protocols, 48-49
RTP traffic, 49
overview, 43-45
classifier tool, 32
classify and police models, 958-963
cloud services, 120
color-aware policing, 73
compression, 172-173
compression strategies over VPN
cRTP and IPsec incompatibilities, 887
overview, 885
TCP optimization using WAAS (wide area application services), 885-886
voice codecs over VPN connection, using, 886-887

conditional trust, 228-230
conditional trust models
campus access (Cisco Catalyst 3750) QoS design, 253-254
campus core (Cisco Catalyst 6500) 
QoS design, 330-332
Cisco Catalyst 4500, 290-291
conditionally trusted endpoints, 230
configuration
Cisco Catalyst 4500, 277
data center virtual access (Nexus 1000V) QoS design, 539-540
FNF (Flexible NetFlow)
  flow exporter, configuring, 149-150
  flow monitor, configuring, 151-152
  flow record, configuring, 150-151
  interface, enabling FNF on relevant, 152
  overview, 149
Mediatrace, 123
Performance Monitor, 125-127
congestion avoidance
described, 85
recommendations and guidelines, 95-96
tools for
  overview, 92
  RED (random early detection), 93
  WRED (weighted random early detection), 93-95
congestion management
overview, 84-85
queuing, levels of, 85-86
queuing tools
  class-based queuing (policy maps), 86-90
  overview, 86
  Tx-Ring operation, 91
recommendations and guidelines, 95-96
scheduling algorithms, 85
terminology, 84
tools, 7
congestion notification, 515-516
contention window (CW), 378-382
Control and Wireless Access Points (CAPWAP), 40, 389
control CBWFQs, 691
control plane policing. See CPP/CoPP (control plane policing)
control plane traffic
  network control, 181
  OAM (operations/administration/management), 182
  overview, 180
  signaling, 181
converged access (Cisco Catalyst 3850 and Cisco 5760 Wireless LAN controller) QoS design
Cisco Catalyst 3850 QoS architecture, 439-442
converged access, 438
enabling QoS, 442-444
ingress QoS models
  classification, marking, and policing model, 448-454
  classification and marking model, 446-448
  overview, 444
  wired-only conditional trust model, 444-446
overview, 435-438
queuing models
  overview, 454
  wired 1P7Q3T egress queuing model, 456-459
wired 2P6Q3T egress queuing model, 459-470
wired queuing, 455
wireless 2P2Q egress queuing model, 472-474
wireless queuing, 470-472
SSID-level traffic, 440-441
steps for, 442-474
converged access QoS design (Tifosi Software Inc. case study)
access-edge design for Cisco IP phones and PCs (conditional trust and classification and marking), 482-485
access-edge design for Cisco TelePresence endpoints (conditional trust), 482
access-edge design for mobile wireless clients (dynamic policy with classification and marking), 489-490
access-edge design for wired access endpoints (DSCP trust), 481-482
access-edge design for wired printer endpoints (no trust), 481
access-edge wired queuing design, 485-488
access-edge wireless queuing design, 491-492
Cisco ISE (Identity Services Engine), 495
CT 5760 Wireless LAN controller uplink ports, 493-495
overview, 477-479
SSID bandwidth allocation between guest and enterprise SSIDs (SSID policy to separate bandwidth distribution), 492-493
wired policies, 481-488
wireless policies, 488-495
core layer (40GE) core-link design, 368-370
core layer (10GE) downlink design, 364-368
core layer Nexus 7000 QoS design, 666-672
CoS (class of service), 32, 572-573
CoS 3 overlap considerations and tactical options, 523-525
CoS/DSCP marking model, 523
CPP/CoPP (control plane policing)
branch router (Cisco ISR G2) QoS design, 757
campus core (Cisco Catalyst 6500) QoS design, 344
campus QoS design considerations and recommendations, 243-244
Cisco Catalyst 3850, 987-990
Cisco Catalyst 4500, 303, 989-996
Cisco Catalyst 6500, 996-998
data center core (Nexus 7000) QoS design, 648
deploying, 987-990
IOS control plane policing, 998-1001
overview, 74-75, 983-985
recommendations, 208-209
traffic classes, defining, 985-987
WAN aggregator (Cisco ASR 1000) QoS design, 708, 733
WAN and branch QoS design considerations and recommendations, 687, 692
CQ (custom queuing), 86
cRTP and IPsec incompatibilities, 887
CS (class selector), 33
CSMA/CA (carrier sense multiple access with collision avoidance), 377-378
CSMA/CD (carrier sense multiple access with collision detection), 377-378
CT 5760 Wireless LAN controller uplink ports, 493-495
CUCM (Cisco Unified Communications Manager), 103
custom protocol classification, 752-753
CUWN (Cisco Unified Wireless Networking), 435-436
CW (contention window), 378-382
\( C_W^{\text{max}} \), 386-387
\( C_W^{\text{min}} \), 386-387

data applications
best effort data, 179
bulk data (high-throughput data), 178-179
overview, 177-178
scavenger (lower-priority data), 180
transactional data (low-latency data), 178
data center access/aggregation (Nexus 5500/2000) QoS design architecture
overview, 562-564
QoS groups and system classes, 567-569
QoS policies supported by, 562-564
VOQ (virtual output queuing), 564-567
egress queuing models
\textit{eight-class model}, 587-591
\textit{four-class model}, 582-587
overview, 582

ingress QoS models
\textit{application policing server models}, 578-580
\textit{buffer size, modifying ingress}, 580-582
\textit{classification and marking models}, 572-578
overview, 569
\textit{trust models}, 570-572
L3 configuration, 592-593
network-qos policy used to set MTU, 597
Nexus 2000 fabric extender QoS, 593-596
overview, 561-562
steps for, 569
data center application-based marking models, 526-527
data center application/tenant-based marking models, 527-528
data center bridging toolset, 508-517
data center core (Nexus 7000) QoS design
additional design options, 638-648
CoPP design, 648
DSCP-mutation model, 645-648
F2/F2e-Series I/O modules, 601, 623-638
fabric modules, 600
M2-Series I/O modules, 601, 604-623
multi-application server classification and marking model, 642-643
overview, 599-604
QoS policies supported by, 601-602
server policing model, 643-645
single-application server marking model, 642
supervisor modules, 600
trust default behavior, 602-603
trusted server model, 638
untrusted server model, 638-642

**data center QoS design (Tifosi Software Inc. case study)**

access/aggregation layer Nexus 5500/2000 QoS design, 659-666
core layer Nexus 7000 QoS design, 666-672
DSCP mutation for signaling traffic between campus and data center, 671-672
multi-application server, 661-662
multi-application virtual machines, 656-657

**network-edge queuing**

*F2 modules*, 666-668
*M2 modules*, 668-671
overview, 657-659, 663-666
overview, 651-655
single-application server, 660-661
single-application virtual machines, 655-656
trusted server, 660
trusted virtual machines, 655
virtual access layer Nexus 1000V QoS design, 655-659

**data center QoS design considerations and recommendations**

architectures

*big data (HPC/HTC/Grid) architectures*, 501-502
*high-performance trading data center architectures*, 500-501
*massively scalable data center architectures*, 506
overview, 500

*secure multitenant data center architectures*, 505

**virtualized multiservice data center architectures**, 503-505
Nexus OS QoS framework, 519-520
overview, 499-500
port QoS roles, 529-531
QoS models

*data center marking models*, 520-528
overview, 520, 528-529
QoS tools

*data center bridging toolset*, 508-517
*data center transmission control protocol (DCTCP)*, 517-519
overview, 507-508

**data center transmission control protocol (DCTCP)**, 517-519

**data center virtual access (Nexus 1000V) QoS design**

architecture, 537-539
configuration notes, 539-540
egress QoS models

*eight-class queuing model*, 556-558
*four-class queuing model*, 551-556
overview, 549-551
ingress QoS models

*classification and marking*, 544-547
overview, 541
*server policing model*, 547-549
trusted models, 541-544
overview, 535-537
statistics, monitoring QoS, 540
trust models

*trusted server model*, 541
untrusted server model, 541-544
VEM (virtual ethernet module), 537-539
VSM (virtual supervisor module), 537-539
data plane policing recommendations, 210-212
data traffic, 4
DBL (dynamic buffer limiting), 278
DC/Campus DSCP mutation, 523
DCBX (data center bridging exchange), 516-517
DCF (Distributed Coordination Function), 376-382
default/best effort CBWFQs, 691
default queuing models
Nexus 7000 (F2/F2e-Series I/O modules), 631-633
Nexus 7000 (M2-Series I/O modules), 608-610
defferential queues, 690
definition of policies (policy maps), 20-22
delay (or latency), 4
deployment principles, 13-14
design principles and strategies
best practice design principles
classification and marking best practices, 191-192
hardware versus software QoS best practices, 190
overview, 189-190
policing and markdown best practices, 192
queuing and dropping best practices, 192-197
QoS design strategies
application class expansion
QoS strategies, 204-205
eight-class model QoS strategy, 200-201
four-class model QoS strategy, 198-199
overview, 198
security QoS strategies, 206-212
twelve-class model QoS strategy, 202-204
security
CPP/CoPP (control plane policing) recommendations, 208-209
data plane policing recommendations, 210-212
overview, 206-208
DF queue recommendations, 195
differentiated services code points. See DSCPs
DiffServ (differentiated services), 6-7, 99. See also specific DiffServ tools
DIFS (DCF interframe space), 378-380
digital signage, 165
distribution layer distribution-link/core-uplink ports, 362-364
distribution layer downlink ports, 360-362
distribution QoS design, 360-364
DMVPN QoS design
challenges, 898-899
example, 900-917
GET VPN
combining, 940
compared, 922-923
hub routers configured for per-tunnel QoS, 901-910
NHRP (next-hop routing protocol), 897-898
overview, 893-894
per-tunnel QoS, 899, 918
role of QoS in a DMVPN network
DMVPN building blocks, 895
overview, 895
where QoS is implemented in DMVPN, 895-896
spoke routers configured for per-tunnel QoS, 910-913
steps for, 901-913
verifying your configuration, 913-917
DoS (denial-of-service) attacks
overview, 34, 206-208
slamming attacks, 206
spoofing attacks, 206
downstream traffic
defining flow, 390
QoS marking strategy, 394-395
wireless LAN controller (Cisco 5500) QoS design, 425-429
DSCPs (differentiated services code points), 6
defined, 33
DSCP-mutation model, 645-648
internal DSCP, 226-227
markings, 191-192
Tifosi Software Inc. (case study), 671-672
transparency, 231
trust DSCP, 228
dual-rate three-color policers, 66-67
dynamic buffer limiting (DBL), 278
dynamic multipoint VPN. See DMVPN QoS design
E-LAN, 774
E-Line, 774
E-Tree, 774
ECN (explicit congestion notification), 685
EDCA (Enhanced Distributed Channel Access)
wireless LAN controller (Cisco 5500) QoS design, 411-412
wireless LAN QoS considerations and recommendations, 382-388
EF queue recommendations: the 33% LLQ rule, 193-195
EFC (ethernet flow control), 508-509
gress QoS models, 238-239
branch router (Cisco ISR G2) QoS design
eight-class model, 754
four-class model, 754
overview, 753
twelve-class model, 754-756
campus access (Cisco Catalyst 3750) QoS design, 265-271
data center virtual access (Nexus 1000V) QoS design
eight-class queuing model, 556-558
four-class queuing model, 551-556
overview, 549-551
enterprise customer edge (Cisco ASR 1000 and ISR G2) QoS design
enterprise-to-service provider mapping models, 798-808
overview, 795
sub-line-rate Ethernet: hierarchical shaping and queuing models, 795-798
enterprise-to-service provider mapping models
eight-class enterprise model mapped to a four-CoS service provider model, 800-803
four-class enterprise model mapped to a four-CoS service provider model, 798-800
overview, 798
twelve-class enterprise model mapped to a four-CoS service provider model, 803-808
sub-line-rate Ethernet: hierarchical shaping and queuing models
known SP policing Bc, 796-797
overview, 795
unknown SP policing Bc, 797-798
WAN aggregator (Cisco ASR 1000) QoS design
eight-class model, 712-715
four-class model, 709-712
overview, 697, 701, 706, 709, 725-726
twelve-class model, 715-725
WAN and branch QoS design considerations and recommendations, 689-692
eight-class queuing models
campus QoS design (Tifosi Software Inc. case study)
eight-class 1P3Q3T egress queuing design, 357-359
eight-class 1P1Q3T ingress queuing design, 355-357
Cisco Catalyst 4500, 281-284
Cisco Catalyst 6500, 314-318
data center access/aggregation (Nexus 5500/2000) QoS design, 587-591
data center virtual access (Nexus 1000V) QoS design, 556-558
GET VPN QoS design, 933-934
Nexus 7000 (F2/F2e-Series I/O modules), 634-637
Nexus 7000 (M2-Series I/O modules), 615-621
eight-CoS fabric QoS policy, 857-858
eight-CoS interface QoS policy, 858-860
eight-CoS SP core model, 857-860
802.11 standard, 35, 374, 382-388
embedded service processors (ESPs), 698-699
endpoints, 119
conditionally trusted endpoints, 230
trusted endpoints, 231
untrusted endpoints, 231
enforcement points, 398
Enhanced Distributed Channel Access. See EDCA
enterprise customer edge (Cisco ASR 1000 and ISR G2) QoS design
egress QoS models
enterprise-to-service provider mapping models, 798-808
overview, 795
sub-line-rate Ethernet: hierarchical shaping and queuing models, 795-798
ingress QoS models, 795
overview, 793
steps for, 794-795
sub-line-rate Ethernet: hierarchical shaping and queuing models, 795-798

enterprise-to-service provider mapping models

- eight-class enterprise model mapped to a four-CoS service provider model, 800-803
- four-class enterprise model mapped to a four-CoS service provider model, 798-800
- overview, 798
- twelve-class enterprise model mapped to a four-CoS service provider model, 803-808

MPLS VPN QoS design considerations and recommendations

- mapping control and signaling traffic, 786
- mapping real-time voice and video, 785-786
- overview, 785
- re-marking and restoring markings, 787
- separating TCP from UDP, 786-787

EPL (ethernet private line), 773

ESPs (embedded service processors), 698-699

EtherChannel QoS design, 234-235

- campus access (Cisco Catalyst 3750) QoS design, 273
- campus core (Cisco Catalyst 6500) QoS design, 343-344
- Cisco Catalyst 4500, 299-300

Ethernet 802.11 WiFi, 38

Ethernet 802.1Q/p, 37

ETS (enhanced transmission selection), 514-515

evolution of QoS, 4-5

EVPL (ethernet virtual private line), 774

explicit congestion notification (ECN), 685

F

FC (priority flow control), 510-512

feature sequencing, 15-16, 52

field values and interpretation, 35-37

FIFO (first-in, first-out), 86

flow-based QoS design, 301-303

flow exporter, configuring, 149-150

flow monitor, configuring, 151-152

flow record, configuring, 150-151

FNF (Flexible NetFlow), 139, 301

AVC (application visibility control) configuration, 149-152

key fields, 148-149

non-key fields, 148-149

overview, 147-148

performance considerations, 159-160

configuration

- flow exporter, configuring, 149-150
- flow monitor, configuring, 151-152
- flow record, configuring, 150-151

interface, enabling FNF on relevant, 152

overview, 149

overview, 147-148
four-class queuing models
Cisco Catalyst 4500, 278-281
Cisco Catalyst 6500, 311-314
data center access/aggregation (Nexus 5500/2000) QoS design, 582-587
data center virtual access (Nexus 1000V) QoS design, 551-556
GET VPN QoS design, 932-933
Nexus 7000 (F2/F2e-Series I/O modules), 634
Nexus 7000 (M2-Series I/O modules), 610-615
four-CoS fabric QoS policy, 850-853
four-CoS interface QoS policy, 853-854
four-CoS SP model, 850-854
frame relay traffic shaping, 78-79

G

GDOI (group domain of interpretation), 923
GET VPN QoS design
building blocks, 924-925
configuration
confirming QoS policy, 936-939
eight-class model, 933-934
four-class model, 932-933
GM (group member) routers, 930-931
KS (key server) routers, 928-929
overview, 931-932
QoS preclassify feature, using, 939-940
twelve-class model, 934-936
DMVPN
combining, 940
compared, 922-923

GDOI (group domain of interpretation), 923
GM (group member) routers, 924-925
IP header preservation, 926-928
KS (key server) routers, 924-925
overview, 921-923, 931-932
service provider, working with, 941
global trends in networking, 164
GM (group member) routers, 924-925, 930-931
Gold QoS profile for wireless LAN controller (Cisco 5500) QoS design, 400-408
GRE handling of MTU issues, 881
Group Encrypted Transport VPN. See GET VPN QoS design
guaranteed-bandwidth queues, 690
guest QoS profile, building, 408-410
guidelines. See recommendations and guidelines

H

hardware
IOS software compared, 678
software QoS best practices compared, 190
headend router configuration, 946-948
hierarchical class-based shaping, 77
hierarchical policing, 23-25, 71
high-definition media, 169
high-definition VoD, 169
high-level packet feature sequence, 16
high-performance trading data center architectures, 500-501
history and evolution
of network infrastructure, 2-5
of packet-switched networks, 3
home office router (spoke) configuration, 948-952
home office VPN (Tifosi Software Inc. case study)
application requirements, 944-945
headend router configuration, 946-948
home office router (spoke) configuration, 948-952
overview, 943-944
QoS configuration, 945-952
HPC/HTC/Grid architectures, 501-502
HPT (high-performance trading) data center architectures, 500-501
HQF (hierarchical queuing framework), 25
HQoS (hierarchical QoS), 776
HTTP sessions, 136
hub routers configured for per-tunnel QoS, 901-910

IETF (Internet Engineering Task Force), 2
ingress QoS models, 235-237
branch router (Cisco ISR G2) QoS design
Medianet classification models, 738-744
NBAR2 classification models, 744-753
overview, 738
campus access (Cisco Catalyst 3750) QoS design, 250-259, 261-265
classification, marking, and policing models, 256-259
classification and marking models, 254-256
converged access (Cisco Catalyst 3850 and Cisco 5760 Wireless LAN controller) QoS design
classification, marking, and policing model, 448-454
classification and marking model, 446-448
overview, 444
wired-only conditional trust model, 444-446
data center access/aggregation (Nexus 5500/2000) QoS design
application policing server models, 578-580
buffer size, modifying ingress, 580-582
classification and marking models, 572-578
overview, 569
trust models, 570-572
data center virtual access (Nexus 1000V) QoS design
classification and marking, 544-547
overview, 541
server policing model, 547-549
trusted models, 541-544
enterprise customer edge (Cisco ASR 1000 and ISR G2) QoS design, 795
Medianet classification models
application-based classification and marking model, 739-743
application-group-based classification model, 743-744
attribute-based classification model, 744
overview, 738-739
NBAR2 classification models
application-based classification and marking model, 745-747
application-group-based classification model, 748
attribute-based classification model, 748-752
custom protocol classification, 752-753
overview, 744-745
overview, 250
trust models
conditional trust models, 253-254
overview, 251
trust CoS model, 251-252
trust DSCP model, 252
untrusted model, 251
WAN aggregator (Cisco ASR 1000) QoS design, 708, 733
WAN and branch QoS design considerations and recommendations, 689
Insight Reporter, 153
interactive video, 34, 164, 166
internal DSCP, 226-227
internal PLIM QoS for ASR 1000, 762-763
Internet edge and AVC (application visibility control), 137, 156-158
Internet Engineering Task Force (IETF), 2
IntServ (integrated services), 6-7
IntServ/DiffServ model
advanced RSVP design, 105-106
basic RSVP design, 104-105
IOS control plane policing, 998-1001
IOS preclassify feature, 877-880
IOS software, 678
IP header preservation, 926-928
IPP (IP precedence), 6
IPsec handling of MTU issues, 881-882
IPsec VPN QoS considerations and recommendations
antireplay implications, 888-890
classification of IPsec packets, 875-876
compression strategies over VPN
cRTP and IPsec incompatibilities, 887
overview, 885
TCP optimization using WAAS (wide area application services), 885-886
voice codecs over VPN connection, using, 886-887
 IOS preclassify feature, 877-880
MTU considerations
GRE handling of MTU issues, 881
IPsec handling of MTU issues, 881-882
overview, 880-881
TCP adjust-MSS feature, 883-885
overview, 871
topologies
IPsec with GRE, 873-874
overview, 871-872
remote-access VPNs, 874-875
standard IPsec VPNs, 872-873
IPSLA Video Operation, 127
IPv4
overview, 39
mapping QoS markings

 Packet classification, 113
 Packet headers, 8, 112
 Packet marking, 114

 IPv6
 Overview, 39, 111-112
 Packet classification, 113
 Packet dropping, 115
 Packet headers, 8, 112
 Packet marking, 114-115
 Policing, 115
 QoS feature support for, 112
 Queuing, 115
 Recommendations and guidelines, 115-116
 Shaping, 115
 Tunneling traffic, 114-115
 ISO (International Organization for Standardization), 3
 ISR G2 routers, 137

 J
 Jitter (or delay variation), 4, 675, 681
 Jitter buffers, 170

 K
 Known SP policing Bc, 796-797
 KS (key server) routers, 924-925, 928-929

 L
 L2 to L3 markings, mapping, 41-42
 L2 tunnels, 40
 L3 tunnels, 40
 LAN-edge QoS policies, 763-765
 Latency, 170
 Propagation, 680-681
 Queuing delay, 681
 Serialization, 680
 WAN and branch QoS design considerations and recommendations, 679-681
 Legacy CLI commands, 25-26
 Link-specific QoS tools, 7
 Link types and speeds, 687-688
 LLQ (low-latency queuing), 87-90
 Policing as part of, 73-74
 RSVP and, 106-107
 WAN and branch QoS design considerations and recommendations, 684
 Load balancing, 234
 Logical or physical interface (class maps), 46
 Lossless transport model
 Data center QoS models, 529
 Port QoS roles, 531

 M
 MAC (media access control), 4
 MAN/WAN Ethernet service evolution, 773-774
 Management and reporting (AVC)
 Insight Reporter, 153
 Overview, 152-153
 Mapping control and signaling traffic, 786
 Mapping QoS markings
 Cisco to RFC 4594 markings, mapping, 42
 L2 to L3 markings, mapping, 41-42
overview, 41
wireless networks, mapping markings for, 43
mapping real-time voice and video, 785-786
markdown
    best practices, 192
tools, 7
markers, policers as, 69
marking, 14, 32
marking-based classifications, 19-20
marking fields in different technologies
    ATM, 38-39
    CAPWAP, 40
    Ethernet 802.11 WiFi, 38
    Ethernet 802.1Q/p, 37
    field values and interpretation, 35-37
    FR, 38-39
    IPv4, 39
    IPv6, 39
    L2 tunnels, 40
    L3 tunnels, 40
    MPLS, 41
    overview, 35
marking tools
    AutoQoS marking, 54
class-based marking (class maps)
        feature sequence, effects of, 52
        overview, 50-52
    packet discard eligibility, 51
    table map feature, mapping markings with, 52-53
    ToS values, 51
    tunnel ToS values, 51
defined, 32
    overview, 50
policing, marking with, 53-54
massively scalable data center architectures, 506
media access control (MAC), 4
media awareness
    flow metadata, 129-130
    MSI (Media Services Interface), 132
    MSP (Media Services Proxy), 132
    NBAR, 130-131
    overview, 121, 127
media content, increase in, 166-167
media monitoring
    IPSLA Video Operation, 127
    Mediatrace
        configuration, 123
        operation, 124-125
        overview, 122-123
    overview, 120, 122
    Performance Monitor
        configuration, 125-127
        overview, 125
Media Session (SIP) snooping, 416-417
Medianet
    architecture and framework, 119-120
    autoconfiguration
        Auto Smartports, 121
        overview, 120-121
    AutoQoS
        Cisco Catalyst 4500 series switches, 971-982
        classify and police models, 958-963
        overview, 121-122, 953-955
        1P3Q3T egress queuing models, 969-971
1P1Q3T ingress queuing models, 968-969
trust models, 955-956
video models, 956-958
VoIP models, 963-968
characteristics of, 118
classification models
application-based classification and marking model, 739-743
application-group-based classification model, 743-744
attribute-based classification model, 744
overview, 738-739
cloud services, 120
endpoints, 119
media awareness
flow metadata, 129-130
MSI (Media Services Interface), 132
MSP (Media Services Proxy), 132
NBAR, 130-131
overview, 121, 127
media monitoring
IPSLA Video Operation, 127
Mediatrace, 122-125
overview, 120, 122
Performance Monitor, 125-127
Medianet metadata classification model, 292-293
network services, 120
overview, 117-119
WAN and branch QoS design considerations and recommendations, 686

Mediatrace
configuration, 123

operation, 124-125
overview, 122-123
MEF (Metro Ethernet Forum), 773
metadata classification, 50
mGRE, 895
microflow policing, 341-342
MLS versus MQC, 225-226
modular QoS command-line framework. See MQC

MPLS, 41
MPLS VPN QoS design
enterprise-to-service provider mapping
mapping control and signaling traffic, 786
mapping real-time voice and video, 785-786
overview, 785
re-marking and restoring markings, 787
separating TCP from UDP, 786-787
MAN/WAN Ethernet service evolution, 773-774
MPLS DiffServ tunneling modes
overview, 781
Pipe Mode, 784-785
Short Pipe Mode, 783-784
Uniform Mode, 782
MPLS VPN architectures, 772
MPLS VPN QoS roles, 787-789
overview, 771-772
QoS paradigm shift, 779-780
service provider class of service models, 781
sub-line-rate Ethernet design implications, 775-778
Tifosi Software Inc. (case study)

*CE router internal QoS (Cisco ASR 1000)*, 863

*CE router LAN-edge QoS policies*, 863

*CE router VPN-edge QoS policies*, 863-866

**overview**, 861-862

*P router interface QoS*, 868

*P router internal QoS (Cisco CRS-3)*, 868

*PE router core-edge QoS*, 867-868

*PE router customer-edge QoS*, 866-867

*PE router internal QoS (Cisco ASR 9000)*, 866

**MQC** classification, 144-147

**MQC (modular QoS command-line) framework**

attaching policies to traffic flows (service policy), 22-23

default behaviors, 19

definition of policies (policy maps), 20-22

hierarchical policies, 23-25

legacy CLI commands, 25-26

**overview**, 16

**syntax**, 17-19

traffic classification (class maps), 19-20

**MSDC** (massively scalable data center) architectures, 506

**MSI** (media services interface), 132

**MSP** (media services proxy), 132

**MTU considerations**

GRE handling of MTU issues, 881

IPsec handling of MTU issues, 881-882

**overview**, 880-881

TCP adjust-MSS feature, 883-885

**multi-action policing**, 71

**multi-application server model**

data center access/aggregation (Nexus 5500/2000) QoS design, 576-578

data center core (Nexus 7000) QoS design, 642-643

data center QoS models, 529

data center virtual access (Nexus 1000V) QoS design, 545-547

port QoS roles, 531

Tifosi Software Inc. (case study), 661-662

**multi-application virtual machines**, 656-657

**multimedia applications**

convergence of media subcomponents within, 168-169

multimedia conferencing, 176-177

multimedia streaming, 177

**overview**, 175-176

**multimedia conferencing**, 34, 176-177

**multimedia/data CBWFQs**, 691

**multimedia streaming**, 34, 177

Multiprotocol Label Switching (MPLS) virtual private network (VPN). See MPLS VPN

N

**NBAR** (network based application recognition), 130-131

metadata classification, 50

**overview**, 47-48

performance routing, 49-50

protocols, 48-49
RTP traffic, 49

**NBAR2 (next generation NBAR)**

AVC (application visibility control)

MQC classification, 144-147

overview, 140-142

performance considerations, 159-160

protocol discovery, 142-144

classification models

application-based classification and marking model, 745-747

application-group-based classification model, 748

attribute-based classification model, 748-752

custom protocol classification, 752-753

overview, 744-745

commands, 115

overview, 140-142

WAN and branch QoS design considerations and recommendations, 687

network control traffic, 181

network downstream, 390

network-edge queuing (Tifosi Software Inc. case study)

F2 modules, 666-668

M2 modules, 668-671

network infrastructure, history and evolution of, 2-5

network-qos policy used to set MTU, 597

network services, 120

network upstream, 390

NHRP (next-hop routing protocol), 895, 897-898

O

OAM (operations/administration/management) traffic, 182

P

P edges, 789

P router interface QoS, 868

P router internal QoS, 868

packet attributes, characteristics, or field values, 45

packet classification

IPv4, 113

IPv6, 113

packet discard eligibility, 51

packet dropping

described, 4

IPv6, 115

packet headers

class-based classification (class maps), 45

IPv4, 112

IPv6, 112

overview, 8

packet jitter. See jitter

packet-loss concealment (PLC), 171

packet marking

IPv4, 114

IPv6, 114-115

packet-switched networks, history and evolution of, 3

partial packets, 777

PDLM (Protocol Description Language Module), 47

PE core-facing edge, 789

PE customer-facing edge, 789
PE ingress/internal QoS (ASR 9000), 789
PE router core-edge QoS, 867-868
PE router customer-edge QoS, 866-867
PE router internal QoS (Cisco ASR 9000), 866
per-port/per-VLAN QoS design, 232-233
campus access (Cisco Catalyst 3750) QoS design, 272-273
Cisco Catalyst 4500, 298-299
per-tunnel QoS between spokes, 918
per-tunnel QoS for DMVPN feature, 899
per-VLAN QoS design
campus access (Cisco Catalyst 3750) QoS design, 271-272
campus core (Cisco Catalyst 6500) QoS design, 342-343
Cisco Catalyst 4500, 297-298
percentage-based policing, 72
percentage-based shaping, 77-78
performance considerations
AVC (application visibility control), 159-160
FNF (Flexible NetFlow), 159-160
NBAR2, 159-160
Performance Monitor, 125-127
performance routing, 49-50
permanent virtual circuit (PVC), 3
PHBs (per-hop behaviors), 6
PINs (places in the network), 2
pipe mode
ingress policer, 827-829
MPLS DiffServ tunneling, 826-834
MPLS EXP-based egress queuing policy, 830-831
MPLS EXP-to-QG ingress mapping policy, 831-832
overview, 784-785
QG-based egress queuing policy, 833-834
platform-specific QoS design options
campus access (Cisco Catalyst 3750) QoS design
AutoQoS SRND4, 274
EtherChannel QoS design, 273
overview, 271
per-port/per-VLAN QoS design, 272-273
per-VLAN QoS design, 271-272
campus core (Cisco Catalyst 6500) QoS design
access-edge design options, 330-340
CPP (control plane policing), 344
EtherChannel QoS design, 343-344
microflow policing, 341-342
overview, 329-330
per-VLAN QoS design, 342-343
Platinum QoS profile for wireless LAN controller (Cisco 5500) QoS design, 400-408
PLC (packet-loss concealment), 171
PLIM (physical layer interface module)
internal PLIM QoS for ASR 1000, 762-763
SIP-based PLIM QoS for ASR 1000, 762
SPA-based PLIM QoS for ASR 1000, 762-763
PoA (point of attachment), 436
policers
best practices, 192

data center virtual access (Nexus 1000V) QoS design, 545-547
defined, 60
dual-rate three-color policers, 66-67
IPv6, 115
as markers, 69
marking with, 53-54

network, placing in, 61
re-mark/markdown, 62

recommendations and guidelines, 79

security and, 68

shapers compared, 60, 777-778
single-rate three-color policers, 65-66
single-rate two-color policers, 64-65
tail drop, 61-62
traffic types, 62
types of, 64-67

policing tools
class-based policing (policy maps)
  color-aware policing, 73
  hierarchical policing, 71
  low-latency queuing, policing as part of, 73-74
  multi-action policing, 71
  overview, 69-70
percentage-based policing, 72
percentage-based shaping, 77-78
PQ (priority queuing), 86
PQ-WFQ (IP RTP priority queuing), 87
WFQ (weighted fair queuing), 87

PoP (point of presence), 436

ports
class maps, 46
QoS roles, 232-233, 529-531

switch ports
  connecting to conditionally trusted endpoints, 240
  connecting to network infrastructure, 241
  connecting to trusted endpoints, 240
  connecting to untrusted endpoints, 240

post-queuing, 15
PQ (priority queuing), 86
PQ-WFQ (IP RTP priority queuing), 87

pre-queuing, 15
principal functions of QoS, 14-15

printer endpoints, access-edge design for, 351

propagation, 680-681
protocols, 48-49
  class-based classification (class maps), 45-46
  data center, 521-523
  NBAR2, 142-144
  storage virtualization, 522-523
PSTN (public switched telephone network), 3
PVC (permanent virtual circuit), 3

Q

QFPs (Quantum Flow Processor), 699-700
QoE, user expectations, 6
QoS (quality of service)
  admission control, 14
  architectural framework, 14-16
  AutoQoS, 25-28
  AVC (application visibility control)
    Internet edge, deploying AVC
    QoS controls at, 156-158
    overview, 154
    WAN edge, deploying AVC
    QoS controls at, 154-156
  bandwidth allocation, 14
  behavioral model, 15
  changes in, 1-2
  classification, 14-15
  classification and marking tools, 7
  congestion management or scheduling tools, 7
  deployment principles, 13-14
  DiffServ (differentiated services), 6-7
  evolution of, 4-5
  feature sequencing, 15-16
  high-level packet feature sequence, 16
  IntServ (integrated services), 6-7
  link-specific tools, 7
  marking, 14
  MQC (modular QoS command-line) framework
    attaching policies to traffic flows (service policy), 22-23
    default behaviors, 19
    definition of policies (policy maps), 20-22
    hierarchical policies, 23-25
    legacy CLI commands, 25-26
    overview, 16
    syntax, 17-19
    traffic classification (class maps), 19-20
  overview, 1-2, 5
  paradigm shift, 779-780
  policing (dropping and markdown), 14
  policing, shaping, and markdown tools, 7
  post-queuing, 15
  pre-queuing, 15
  principal functions of, 14-15
  queuing, 14-15
  shaping, 14
  simplification/automation of, 9
  standardization and consistency, 9-10
  standards evolution
    overview, 183
    RFC 2597 (clarification), 183-184
    RFC 4594 (update draft), 185-187
    RFC 5865 (proposed standard), 184-185
toolset, 7-8
user expectations, 6
QoS preclassify feature, using, 939-940
QoX, 6
Quantum Flow Processor (QFPs), 699-700
queuing, 14-15
best practices
AF queue recommendations, 195
DF queue recommendations, 195
EF queue recommendations: the 33% LLQ rule, 193-195
overview, 192-193
scavenger class queue recommendations, 195-196
WRED recommendations, 197
deferential queues, 690
defined, 84
guaranteed-bandwidth queues, 690
IPv6, 115
levels of, 85-86
real-time queues, 690
WAN and branch QoS design considerations and recommendations, 689-692
queuing delay, 681
queuing models
campus core (Cisco Catalyst 6500) QoS design
eight-class (8Q4T ingress and 1P7Q4T egress) queuing models, 314-318
four-class (4Q4T ingress and 1P3Q4T egress) queuing models, 311-314
overview, 308-311
2P6Q4T ingress and egress queuing models, 328-329
twelve-class (8Q4T ingress and 1P7Q4T egress) queuing models, 318-328
Cisco Catalyst 4500
eight-class egress queuing model, 281-284
four-class egress queuing model, 278-281
overview, 277-281
twelve-class egress queuing model, 284-289
converged access (Cisco Catalyst 3850 and Cisco 5760 Wireless LAN controller) QoS design
overview, 454
wired 1P7Q3T egress queuing model, 456-459
wired 2P6Q3T egress queuing model, 459-470
wired queuing, 455
wireless 2P2Q egress queuing model, 472-474
wireless queuing, 470-472
Nexus 7000 (F2/F2e-Series I/O modules)
default queuing models, 631-633
eight-class (4Q1T ingress/1P3Q1T egress) queuing model, 634-637
four-class (4Q1T ingress/1P3Q1T egress) queuing model, 634
overview, 630
Nexus 7000 (M2-Series I/O modules)
default queuing models, 608-610
eight-class (8Q2T ingress/1P3Q4T egress) queuing model, 615-621
four-class (4Q2T ingress/1P3Q4T egress) queuing model, 610-615 overview, 607
queuing tools
class-based queuing (policy maps)
  CBWFQ (class-based weighted fair queuing), 87-89
  CQ (custom queuing), 86
  FIFO (first-in, first-out), 86
  LLQ (low-latency queuing), 87-90
  overview, 86-87
  PQ (priority queuing), 86
  PQ-WFQ (IP RTP priority queuing), 87
  WFQ (weighted fair queuing), 87
  overview, 86
  Tx-Ring operation, 91
IPv6, 115-116
policing, 79
RSVP, 108
shaping, 79
standards and design guidelines, changes in, 2
RED (random early detection), 93
remote-access VPNs, 874-875
RFC (Request for Comments), 2, 6
  improvements in, 10
  RFC 2597, 183-184
  RFC 3662, 34
  RFC 4594, 10, 182, 185-187
  RFC 4595, 171
  RFC 5865, 10, 184-185
room-based videoconferencing, 166
round-robin queues, 85
RSVP (Resource Reservation Protocol)
  branch router (Cisco ISR G2) QoS design, 757
deployment models
  IntServ/DiffServ model (advanced design), 105-106
  IntServ/DiffServ model (basic design), 104-105
  overview, 103-104
LLQ and, 106-107
  overview, 6, 100-102
proxy, 102-103
recommendations and guidelines
  classification and marking, 55
  congestion avoidance, 95-96
  congestion management, 95-96
  WAN aggregator (Cisco ASR 1000) QoS design
    advanced RSVP model with application ID, 729-733
    basic RSVP model, 726-729
    overview, 697, 701, 706, 709, 725-726
  WAN and branch QoS design considerations and recommendations, 685-686
RTP traffic, 49
radio downstream, 390
radio upstream, 389
random dropping, 62
re-mark/markdown policers, 62
re-marking and restoring markings, 787
real-time interactive video, 34, 174-175
real-time queues, 690
recommendations and guidelines
classification and marking, 55
congestion avoidance, 95-96
congestion management, 95-96
scavenger (lower-priority data), 180
scavenger CBWFQs, 691
scavenger class queue recommendations, 195-196
scheduling algorithms, 85
secure multitenant data center architectures, 505
security
DoS (denial-of-service) attacks, 34
overview, 206-208
slamming attacks, 206
spoofing attacks, 206
network attacks, 34
and policers, 68
QoS design strategies
CPP/CoPP (control plane policing) recommendations, 208-209
data plane policing recommendations, 210-212
overview, 206-208
trust boundaries, 33
worms, 34, 206-208
serialization, 680
server policing model
data center QoS models, 529
port QoS roles, 531
service-policy command, 17-19
service provider, working with, 941
service provider core (Cisco CRS) QoS design
architecture, 846-849
design steps, 849
overview, 845-846
SP core CoS QoS models
eight-CoS SP core model, 857-860
four-CoS SP model, 850-854
overview, 849-850
six-CoS SP core model, 854-857
service provider edge (Cisco ASR 9000) QoS design architecture, 810-814
MPLS DiffServ tunneling models
overview, 814-815
pipe mode MPLS DiffServ tunneling, 826-834
short pipe mode MPLS DiffServ tunneling, 834-842
uniform mode MPLS DiffServ tunneling, 815-826
overview, 809
steps for, 814
service set identifiers. See SSID
shapers
defined, 60
IPv6, 115
network, placing in, 61
overview, 14
partial packets, 777
policers compared, 60, 777-778
recommendations and guidelines, 79
software algorithm to enforce packets to delay, 777
tail drop, 61-62
traffic types, 62
shaping tools
class-based shaping (policy maps)
hierarchical class-based shaping, 77
overview, 76-77
percentage-based shaping, 77-78
legacy shaping tools
  ATM traffic shaping, 78
  frame relay traffic shaping, 78-79
overview, 78
overview, 75-76
QoS, 7
short pipe mode
  DSCP-based egress queuing policy, 840-842
  ingress policer, 835-838
  MPLS DiffServ tunneling, 834-842
  MPLS EXP-based egress queuing policy, 838-840
overview, 783-784
signaling, 181
Silver QoS profile for wireless LAN controller (Cisco 5500) QoS design, 400-408
simplification/automation of QoS, 9
single-application server, 660-661
  data center access/aggregation (Nexus 5500/2000) QoS design, 573-576
  data center QoS models, 528
  data center virtual access (Nexus 1000V) QoS design, 544-545
  port QoS roles, 530
single-application virtual machines, 655-656
single-rate three-color policers, 65-66
single-rate two-color policers, 64-65
SIP-based PLIM, 762
SIP-10s oversubscription scenarios, six-CoS fabric QoS policy, 855-856
six-CoS interface QoS policy, 856-857
six-CoS SP core model, 854-857
skid buffers, 512-514
slamming attacks, 206
smartphones, use of, 167
SMDC (secure multitenant data center) architectures, 505
SNA (Systems Network Architecture), 3
social networking, appearance and effect on business networks of, 167
software algorithm to enforce packets to delay, 777
SP core CoS QoS models
  eight-CoS SP core model, 857-860
  four-CoS SP model, 850-854
  overview, 849-850
  six-CoS SP core model, 854-857
SPA-based matrix of ingress classification by SIP or SPA level, 705-706
SPA-based PLIM, 762-763
Spectralink voice priority, 411
SPGs (switch peer groups), 436
spoke routers configured for per-tunnel QoS, 910-913
spoofing attacks, 206
SSID (service set identifier)
  overview, 35
  SSID bandwidth allocation between guest and enterprise SSIDs (SSID policy to separate bandwidth distribution), 492-493
  SSID-level traffic, 440-441
standard IPsec VPNs, 872-873
standardization and consistency, 9-10
standards and design guidelines, changes in, 2
statistics, monitoring QoS, 540
storage virtualization protocols, 522-523
strategic QoS design (Tifosi Software Inc. case study)
business catalysts for QoS reengineering, 216-217
eight-class QoS model, challenges, 219-220
eight-class QoS model, proposed, 217-219
four-class QoS model, original, 215-216
overview, 215
streaming video, 34, 164-165
strict priority queues, 85
sub-line-rate Ethernet: hierarchical shaping and queuing models
known SP policing Bc, 796-797
overview, 795
unknown SP policing Bc, 797-798
sub-line-rate Ethernet design implications, 775-778
SVC (switched virtual circuit), 3
switch peer groups (SPGs), 436
switch ports
connecting to conditionally trusted endpoints, 240
connecting to network infrastructure, 241
connecting to trusted endpoints, 240
connecting to untrusted endpoints, 240
syntax for MQC (modular QoS command-line) framework, 17-19
Systems Network Architecture (SNA), 3

T

table map feature, mapping markings with, 52-53
tail drop policers/shapers, 61-62
TCP adjust-MSS feature, 883-885
TCP optimization using WAAS (wide area application services), 885-886
template generation and installation (AutoQoS), 28
terminology
classification and marking, 32-33
congestion management, 84
TID (traffic identifier), 33
Tifosi Software Inc. (case study)
campus QoS design
access layer uplink design, 359-360
access QoS design, 350-360
Cisco Catalyst 3750, 350-360
Cisco Catalyst 4550, 360-364
Cisco Catalyst 6550, 364-370
Cisco IP phones or PCs (conditional trust and classification and marking), access-edge design for, 352-355
Cisco TelePresence endpoints (conditional trust), access-edge design for, 352
core layer (40GE) core-link design, 368-370
core layer (10GE) downlink design, 364-368
core QoS design, 364-370
distribution layer distribution-link/core-uplink ports, 362-364
distribution layer downlink ports, 360-362
distribution QoS design, 360-364
eight-class 1P3Q3T egress queuing design, 357-359
eight-class 1P1Q3T ingress queuing design, 355-357
overview, 347-350
printer endpoints, access-edge design for, 351
wireless access endpoints (DSCP Trust), access-edge design for, 351
converged access QoS design
access-edge design for Cisco IP phones and PCs (conditional trust and classification and marking), 482-485
access-edge design for Cisco TelePresence endpoints (conditional trust), 482
access-edge design for mobile wireless clients (dynamic policy with classification and marking), 489-490
access-edge design for wired access endpoints (DSCP trust), 481-482
access-edge design for wired printer endpoints (no trust), 481
access-edge wired queuing design, 485-488
access-edge wireless queuing design, 491-492
Cisco ISE (Identity Services Engine), 495
CT 5760 Wireless LAN controller uplink ports, 493-495
overview, 477-479
SSID bandwidth allocation between guest and enterprise SSIDs (SSID policy to separate bandwidth distribution), 492-493
wired policies, 481-488
wireless policies, 488-495
data center QoS design
access/aggregation layer Nexus 5500/2000 QoS design, 659-666
core layer Nexus 7000 QoS design, 666-672
DSCP mutation for signaling traffic between campus and data center, 671-672
multi-application server, 661-662
multi-application virtual machines, 656-657
network-edge queuing, 657-659, 663-666
network-edge queuing (F2 modules), 666-668
network-edge queuing (M2 modules), 668-671
overview, 651-655
single-application server, 660-661
single-application virtual machines, 655-656
trusted server, 660
trusted virtual machines, 655
virtual access layer Nexus 1000V QoS design, 655-659
home office VPN
application requirements, 944-945
overview, 943-944
QoS configuration, 945-952
MPLS VPN QoS design
- CE router internal QoS (Cisco ASR 1000), 863
- CE router LAN-edge QoS policies, 863
- CE router VPN-edge QoS policies, 863-866
  overview, 861-862
- P router interface QoS, 868
- P router internal QoS (Cisco CRS-3), 868
- PE router core-edge QoS, 867-868
- PE router customer-edge QoS, 866-867
- PE router internal QoS (Cisco ASR 9000), 866
  overview, 215

strategic QoS design
- business catalysts for QoS reengineering, 216-217
- eight-class QoS model, challenges, 219-220
- eight-class QoS model, proposed, 217-219
- four-class QoS model, original, 215-216

WAN and branch QoS design
- internal PLIM QoS for ASR 1000, 762-763
- LAN-edge QoS policies, 763-765
  overview, 759-760
- WAN-edge QoS policies, 765-768

token bucket algorithms, 62-64

top-down deployments, 168
topologies for IPsec VPN
- IPsec with GRE, 873-874

overview, 871-872
remote-access VPNs, 874-875
standard IPsec VPNs, 872-873
ToS (type of service), 32, 51

traffic classes
- characteristics of, 4
- CPP/CoPP, 985-987
  guidelines for, 10
  overview, 4

traffic identifier (TID), 33

transactional data (low-latency data), 178

trust boundaries, 230-231, 399-400
trust CoS, 228, 251-252
trust DSCP, 228, 252
trust policy, 443-444, 446
trust states and operations, 227-230
trusted endpoints, 231

trusted server models
- data center access/aggregation (Nexus 5500/2000) QoS design, 570
- data center core (Nexus 7000) QoS design, 638
- data center QoS models, 528
- data center virtual access (Nexus 1000V) QoS design
  trusted server model, 541
  untrusted server model, 541-544
  port QoS roles, 530
  Tifosi Software Inc. (case study), 660

trusted virtual machines, 655

TSpec (transmission specification), 388

tunnel ToS values, 51
tunneling traffic, 114-115
twelve-class queuing models
Cisco Catalyst 4500, 284-289
Cisco Catalyst 6500, 318-328
GET VPN QoS design, 934-936
Tx-Ring, 91, 682-683
TXOP (transmission opportunity), 388
type of service (ToS), 32, 51

UDP (User Datagram Protocol), 93
unconditional packet drop, 75
uniform mode
ingress policer, 816-821
MPLS DiffServ tunneling, 815-826
MPLS EXP-based egress queuing policy, 822-823
MPLS EXP-to-QG ingress mapping policy, 823-824
overview, 782
QG-based egress queuing policy, 824-826
unknown SP policing Bc, 797-798
untrusted endpoints, 231
untrusted server model
Cisco Catalyst 3750, 251
data center access/aggregation (Nexus 5500/2000) QoS design, 570-572
data center QoS models, 528
data center virtual access (Nexus 1000V) QoS design, 541-544
Nexus 7000, 638-642
port QoS roles, 530
untrusted state, 227
upstream QoS marking strategy, 392-394
upstream traffic, 389-390, 429-430
user expectations, 6

V
VEM (virtual ethernet module), 537-539
video applications
broadcast video, 173-174
compression, 172-173
evolution of, 164-166
interactive video, 164, 166
optimized priority, 411
overview, 171-173
real-time interactive, 174-175
streaming video, 164-165
video conferencing, 166
video surveillance, 165
video traffic
broadcast video, 34
categories of, 4-5
classification and marking, 34
growth of, 2
interactive video, 34
multimedia conferencing, 34
multimedia streaming, 34
overview, 4
real-time interactive video, 34
streaming video, 34
virtual access layer Nexus 1000V QoS design, 655-659
virtualized multiservice data center architectures, 503-505
VLAN-based QoS, 232-233
VMDC (virtualized multiservice data center) architectures, 503-505
VMs (virtual machines), 535-537. See also data center virtual access (Nexus 1000V) QoS design.

VoD streams, 165

Voice
bandwidth, 171
optimized priority, 411
overview, 170-171
recommendations, 170
requirements, 170
traffic, 4

Voice codecs over VPN connection, using, 886-887

VoIP
AutoQoS, 963-968
Cisco 5500, 410-413
jitter buffers, 170
latency, 170
PLC (packet-loss concealment), 171

VOQs (virtual output queues), 512-514, 564-567, 605
VQIs (virtual queuing indexes), 605
VSM (virtual supervisor module), 537-539

WAN aggregation routers, 677-678
WAN aggregator ingress/internal QoS, 692
WAN aggregator LAN edge, 693
WAN aggregator (Cisco ASR 1000) QoS design
additional platform-specific QoS design options, 725-733
architecture, 698-700
AutoQoS SRND4, 733
ccontrol plane policing, 733
egress QoS models
   eight-class model, 712-715
   four-class model, 709-712
   overview, 709
twelve-class model, 715-725

ESPs (embedded service processors), 698-699

Ingress QoS models
   overview, 701
   SIP-based PLIM, 707-708
   SIP-10s oversubscription scenarios, 703
   SPA-based matrix of ingress classification by SIP or SPA level, 705-706
   SPA-based PLIM, 706-707
   overview, 697, 701, 706, 709, 725-726
   QFPs (Quantum Flow Processor), 699-700

RSVP
   advanced RSVP model with application ID, 729-733
   basic RSVP model, 726-729
   overview, 725-726
   steps for, 700

WAN aggregator WAN edge, 693

WAN and branch QoS design (Tifosi Software Inc. case study)
   internal PLIM QoS for ASR 1000, 762-763
   LAN-edge QoS policies, 763-765
   overview, 759-760
   WAN-edge QoS policies, 765-768

WAN and branch QoS design considerations and recommendations
   architectures, 677
AVC (application visibility control), 687
branch interface QoS roles, 692-693
CBWFQ (class-based weighted fair queuing), 683
CPP (control plane policing), 687
hardware versus IOS software, 678
jitter, 681
latency, 679-681
link types and speeds, 687-688
LLQ (low-latency queuing), 684
Mediant, 686
NBAR2, 687
overview, 675-676
QoS models
  CPP (control plane policing), 692
  egress QoS models, 689-692
  ingress QoS models, 689
  overview, 688-689
RSVP (Resource Reservation Protocol), 685-686
Tx-Ring, 682-683
WRED (weighted random early detect), 685
WAN edge, 137, 154-156, 765-768
WebEx, 118
WFQ (weighted fair queuing), 85, 87
wired and wireless LAN environments compared, 374-376
wired-only conditional trust model, 444-446
wired policies, 481-488
wired 1P7Q3T egress queuing model, 456-459
wired 2P6Q3T egress queuing model, 459-470
wired queuing, 455
wireless access
changes in, 2
endpoints (DSCP Trust), access-edge design for, 351
mapping markings for, 43
overview, 4
WLAN QoS profiles, 400-408
wireless LAN controller (Cisco 5500) QoS design
AVC (application visibility control), 417-424
Bronze QoS profile, 400-408
CAC (call admission control)
  configuring, 414-415
  overview, 413
downstream traffic, 425-429
EDCA, optimizing, 411-412
eight-class model design, 430-431
enforcement points, 398
four-class model design, 425-430
Gold QoS profile, 400-408
guest QoS profile, building, 408-410
Media Session (SIP) snooping, 416-417
overview, 397
Platinum QoS profile, 400-408
Silver QoS profile, 400-408
strategy, developing, 424-431
trust boundaries, 399-400
twelve-class model design, 431
upstream traffic, 429-430
VoIP applications, 410-413
WLAN QoS profiles, 400-408
WMM policy
  enabling, 413-414
  overview, 405-408
wireless LAN QoS considerations and recommendations
ACs (access categories), 383-385
AIFSN (arbitration interframe spacing number), 385-386
building blocks for, 376-382
CAPWAP (Control and Wireless Access Points), 389
CSMA/CD (carrier sense multiple access with collision detection), 377-382
CW (Contention Window), 378-382
\( CW_{\text{max}} \), 386-387
\( CW_{\text{min}} \), 386-387
DCF (Distributed Coordination Function), 376-382
downstream QoS marking strategy, 394-395
downstream traffic flow, defining, 390
EDCA (Enhanced Distributed Channel Access), 382-388
t802.11e standard, 382-388
overview, 373-374, 389
QoS mappings and markings, 390-391
TSpec (transmission specification), 388
TXOP (transmission opportunity), 388
upstream QoS marking strategy, 392-394
upstream traffic flow, defining, 389-390
wired and wireless LAN environments compared, 374-376

wireless policies, 488-495
wireless 2P2Q egress queuing model, 472-474

wireless queuing, 470-472
wireless traffic, 35
WMM (Wireless Multimedia), 374, 438
WMM policy
  enabling, 413-414
  overview, 405-408
worms, 34, 206-208
WRED (weighted random early detection), 93-95, 197, 685

wireless queuing, 470-472
wireless traffic, 35
WMM (Wireless Multimedia), 374, 438
WMM policy
  enabling, 413-414
  overview, 405-408
worms, 34, 206-208
WRED (weighted random early detection), 93-95, 197, 685