CHAPTER 4

Medianet QoS Design Considerations

This document provides an overview of Quality of Service (QoS) tools and design recommendations relating to an enterprise medianet architecture and includes high-level answers to the following:

- Why is Cisco providing new QoS design guidance at this time?
- What is Cisco's Quality of Service toolset?
- How can QoS be optimally deployed for enterprise medianets?

QoS has proven itself a foundational network infrastructure technology required to support the transparent convergence of voice, video, and data networks. Furthermore, QoS has also been proven to complement and strengthen the overall security posture of a network. However, business needs continue to evolve and expand, and as such, place new demands on QoS technologies and designs. This document examines current QoS demands and requirements within an enterprise medianet and presents strategic design recommendations to address these needs.

Drivers for QoS Design Evolution

There are three main sets of drivers pressuring network administrators to reevaluate their current QoS designs (each is discussed in the following sections):

- New applications and business requirements
- New industry guidance and best practices
- New platforms and technologies

New Applications and Business Requirements

Media applications—particularly video-oriented media applications—are exploding over corporate networks, exponentially increasing bandwidth utilization and radically shifting traffic patterns. For example, according to recent studies, global IP traffic will nearly double every two years through 2012[1] and the sum of all forms of video will account for close to 90 percent of consumer traffic by 2012[2].

1. Cisco Visual Networking Index—Forecast and Methodology, 2007-2012
   http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_Whitd_Paper.html
2. Approaching the Zettabyte Era
   http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481374_ns827_Networking_Solutions_White_Paper.html

Businesses recognize the value that media applications—particularly video-based collaborative applications—bring to the enterprise, including:

- Increasing productivity
- Improving the quality of decision making
- Speeding time-to-market
- Facilitating knowledge sharing
- Fueling innovation
- Reducing travel time and expenses
- Protecting the environment

Corresponding to these values and benefits of media applications, there are several business drivers behind media application growth, including:

- Evolution of video applications
- Transition to high-definition media
- Explosion of media
- Phenomena of social networking
- Emergence of “bottoms-up” media applications
- Convergence within media applications
- Globalization of the workforce
- Pressures to go green

These business drivers briefly described in the following sections.

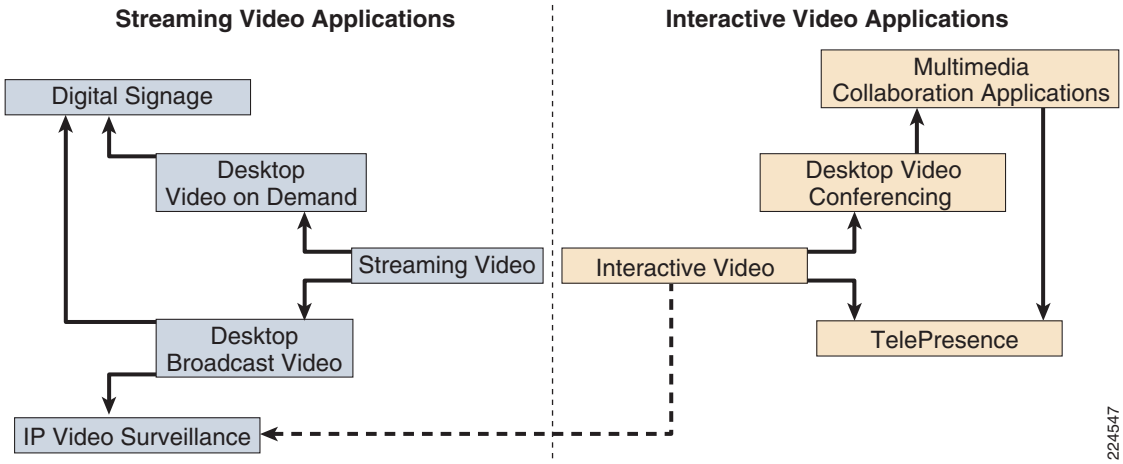
The Evolution of Video Applications

When the previous Cisco Enterprise QoS Design Guide was published (in 2003), there were basically only two broad types of video applications deployed over enterprise networks:

- Interactive video—Generally describes H.323-based collaborative video applications (typically operating at 384 kbps or 768 kbps); video flows were bi-directional and time-sensitive.
- Streaming video—Generally describes streaming or video-on-demand (VoD) applications; video flows were unidirectional (either unicast or multicast) and were not time-sensitive (due to significant application buffering).

However, at the time of writing this document (2009), video applications have evolved considerably, as illustrated in Figure 4-1.

Figure 4-1 Video Application Evolution



Consider first the streaming video branch—the earliest sets of video applications were VoD streams to the desktop. VoD streams can include pre-recorded content such as employee communications, training, e-learning, and social-interaction content. Today, due to the ease of content creation, on-demand content may either be professionally-produced (top-down) or self-produced (bottom-up). It is important to also note that not all VoD content is necessarily business-related, as non-business, entertainment-oriented content is often widely available for on-demand video viewing.

VoD applications soon expanded to include the development and support of "live" or "broadcast" video streams to the desktop. Broadcast streams may include company meetings, special events, internal corporate announcements or similar content. As such, broadcast streaming video content is typically professionally-produced, top-down content.

Thereafter, with the proliferation of flat-screen digital displays, it became increasingly apparent that the desktop is not the only display option for streaming video. Thus, digital signage began to emerge as another streaming video application (for both on-demand and broadcast video streams). Digital signage refers to centrally-managed publishing solutions for delivering digital media to networked displays. For example, Cisco offers a Digital Media Player and an enterprise TV solution that works in conjunction with its Digital Media System to support a comprehensive digital signage solution. Digital signage can be used to broadcast internal information, such as sharing up-to-date schedules and news where people need it most or providing realtime location and directional guidance. Additionally, digital signage is an effective tool for marketing, helping companies to promote products and services directly to customers.

Around the same time that digital signage was being developed, the advantages that IP brought to video were being gradually being applied to the video surveillance market. These advantages include the ability to forward live video streams to local or remote control centers for observation and efficient processing. Cisco offers comprehensive IP surveillance (IPVS) solutions, including IP cameras, hybrid analog-to-digital video gateways (to facilitate transitioning from closed-circuit TV surveillance solutions to IPVS), and IPVS management applications. Interestingly, video surveillance has a unique degree of interactivity not found in any other streaming video application, namely, that of having an observer "interact" with the video stream by sending control information to the transmitting video camera, for instance, to track an event-in-progress.

On the interactive video side of the video application hemisphere, there has also been considerable application evolution. Basic video conferencing applications, which were initially dedicated room-based units, evolved into software-based PC applications. The factors behind this shift from room-based hardware to PC-based software were two-fold:

- The convenience of immediate desktop collaboration (rather than having to book or hunt for an available video-conferencing enabled room).

- The availability of inexpensive Webcams. Desktop video conferencing may be utilized on a one-to-one basis or may support a few participants simultaneously.

Once video conferencing moved to software, a whole new range of communication possibilities opened up, which morphed desktop video conferencing applications into multimedia collaboration applications. Multimedia collaboration applications, including Cisco Unified Personal Communicator (CUPC) and Cisco WebEx, share not only voice and video, but also data applications, such as instant messaging, document and presentation sharing, application sharing, and other integrated multimedia features.

However, not all interactive video migrated to the desktop. Room-based video conferencing solutions continued to evolve and leveraged advances in high-definition video and audio, leading to solutions like Cisco TelePresence. Additionally, application sharing capabilities—borrowed from multimedia conferencing applications—were added to these high-definition room-based video conferencing solutions.

And video application evolution doesn't end here, but will continue to expand and morph over time as new demands and technologies emerge.

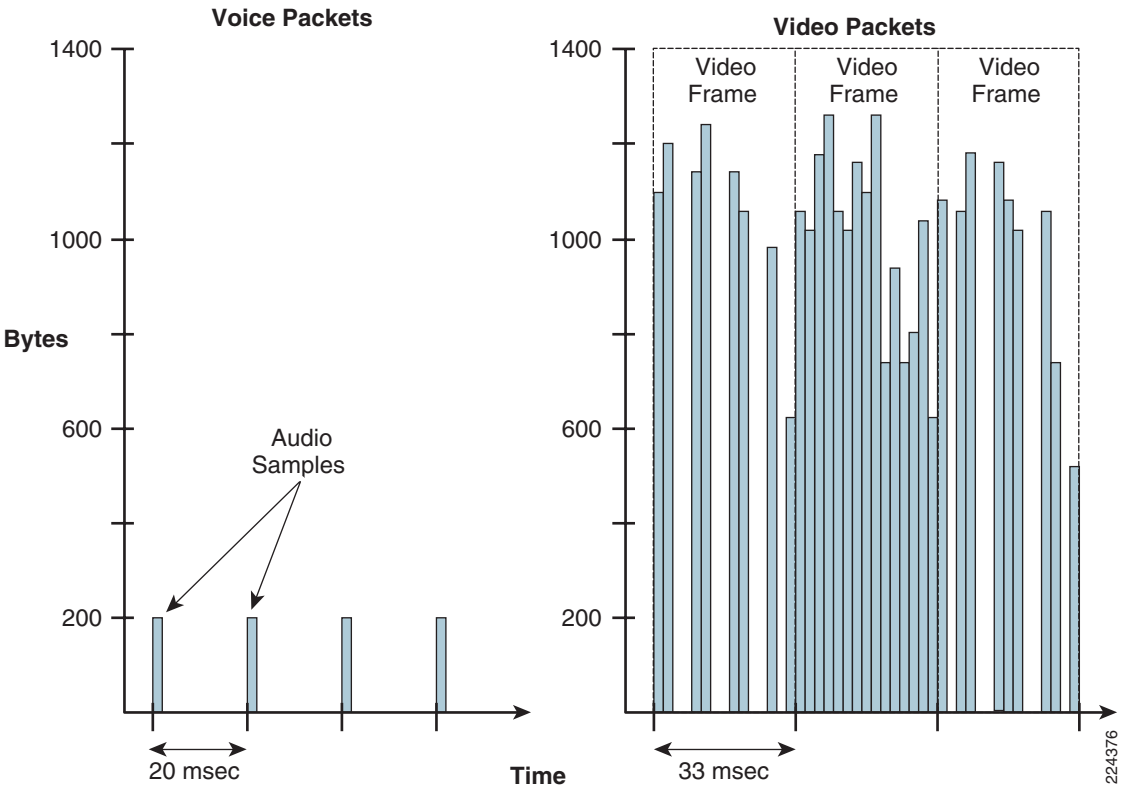
The Transition to High-Definition Media

One of the reasons traditional room-to-room video conferencing and desktop Webcam-style video conferencing are sometimes questioned as less than effective communications systems is the reliance on low-definition audio and video formats.

On the other hand, high-definition interactive media applications, like Cisco TelePresence, demonstrate how high-definition audio and video can create an more effective remote collaboration experience, where meeting participants actually feel like they are in the same meeting room. Additionally, IP video surveillance cameras are migrating to high-definition video in order to have the digital resolutions needed for new functions, such as pattern recognition and intelligent event triggering based on motion and visual characteristics. Cisco fully expects other media applications to migrate to high-definition in the near future, as people become accustomed to the format in their lives as consumers, as well as the experiences starting to appear in the corporate environment.

High-definition media formats transmitted over IP networks create unique challenges and demands on the network that need to be planned for. For example, Figure 4-2 contrasts the behavior of VoIP as compared to high definition video at the packet level.

Figure 4-2 VoIP versus High-Definition Video—At the Packet Level



The network demands of high-definition video include not only radically more bandwidth, but also significantly higher transmission reliability, as compared to standard-definition video applications.

The Explosion of Media

Another factor driving the demand for video on IP networks is the sheer explosion of media content. The barriers to media production, distribution, and viewing have been dramatically lowered. For example, five to ten years ago video cameras became so affordable and prevalent that just about anyone could buy one and become an amateur video producer. Additionally, video cameras are so common that almost every cell phone, PDA, laptop, and digital still camera provides a relatively high-quality video capture capability. However, until recently, it was not that easy to be a distributor of video content, as distribution networks were not common.

Today, social networking sites like YouTube, MySpace and many others appearing every day and have dramatically lowered the barrier to video publishing to the point where anyone can do it. Video editing software is also cheap and easy to use. Add to that a free, global video publishing and distribution system and essentially anyone, anywhere can be a film studio. With little or no training, people are making movie shorts that rival those of dedicated video studios.

The resulting explosion of media content is now the overwhelming majority of consumer network traffic and is quickly crossing over to corporate networks. The bottom line is there are few barriers left to inhibit video communication and so this incredibly effective medium is appearing in new and exciting applications every day.

The Phenomena of Social Networking

Social networking started as a consumer phenomenon, with people producing and sharing rich media communications such as blogs, photos, and videos. When considering the effect it may have on corporate networks, some IT analysts believed social networking would remain a consumer trend, while others believed the appearance in corporate networks was inevitable.

Skeptics look at social networking sites like YouTube, MySpace, and others and see them as fads primarily for the younger population. However, looking beyond the sites themselves, it is important to understand the new forms of communication and information sharing they are enabling. For example, with consumer social networking, typically people are sharing information about themselves, about subjects they have experience in, and interact with others in real-time who have similar interests. In the workplace, we already see parallel activities, because the same types of communication and information sharing are just as effective.

The corporate directory used to consist of employee names, titles, and phone numbers. Companies embracing social networking are adding to that skillsets and experience, URL links to shared work spaces, blogs, and other useful information. The result is a more productive and effective workforce that can adapt and find the skillsets and people needed to accomplish dynamic projects.

Similarly, in the past information was primarily shared via text documents, E-mail, and slide sets. Increasingly, we see employees filming short videos to share best practices with colleagues, provide updates to peers and reports, and provide visibility into projects and initiatives. Why have social networking trends zeroed in on video as the predominant communication medium? Simple: video is the most effective medium. People can show or demonstrate concepts much more effectively and easily using video than any other medium.

Just as a progression occurred from voice exchange to text, to graphics, and to animated slides, video will start to supplant those forms of communications. Think about the time it would take to create a good set of slides describing how to set up or configure a product. Now how much easier would it be just to film someone actually doing it? That's just one of many examples where video is supplanting traditional communication formats.

Internally, Cisco has witnessed the cross-over of such social networking applications into the workplace, with applications like Cisco Vision (C-Vision). C-Vision started as an ad hoc service by several employees, providing a central location for employees to share all forms of media with one another, including audio and video clips. Cisco employees share information on projects, new products, competitive practices, and many other subjects. The service was used by so many employees that Cisco's IT department had to assume ownership and subsequently scaled the service globally within Cisco. The result is a service where employees can become more effective and productive, quickly tapping into each other's experiences and know-how, all through the effectiveness and simplicity of media.

The Emergence of Bottom-Up Media Applications

As demonstrated in the C-Vision example, closely related to the social-networking aspect of media applications is the trend of users driving certain types of media application deployments within the enterprise from the bottom-up (in other words, the user base either demands or just begins to use a given media application with or without formal management or IT support). Bottom-up deployment patterns have been noted for many Web 2.0 and multimedia collaboration applications.

In contrast, company-sponsored video applications are pushed from the top-down (in other words, the management team decides and formally directs the IT department to support a given media application for their user base). Such top-down media applications may include Cisco TelePresence, digital signage, video surveillance, and live broadcast video meetings.

The combination of top-down and bottom-up media application proliferation places a heavy burden on the IT department as it struggles to cope with officially-supported and officially-unsupported, yet highly proliferated, media applications.

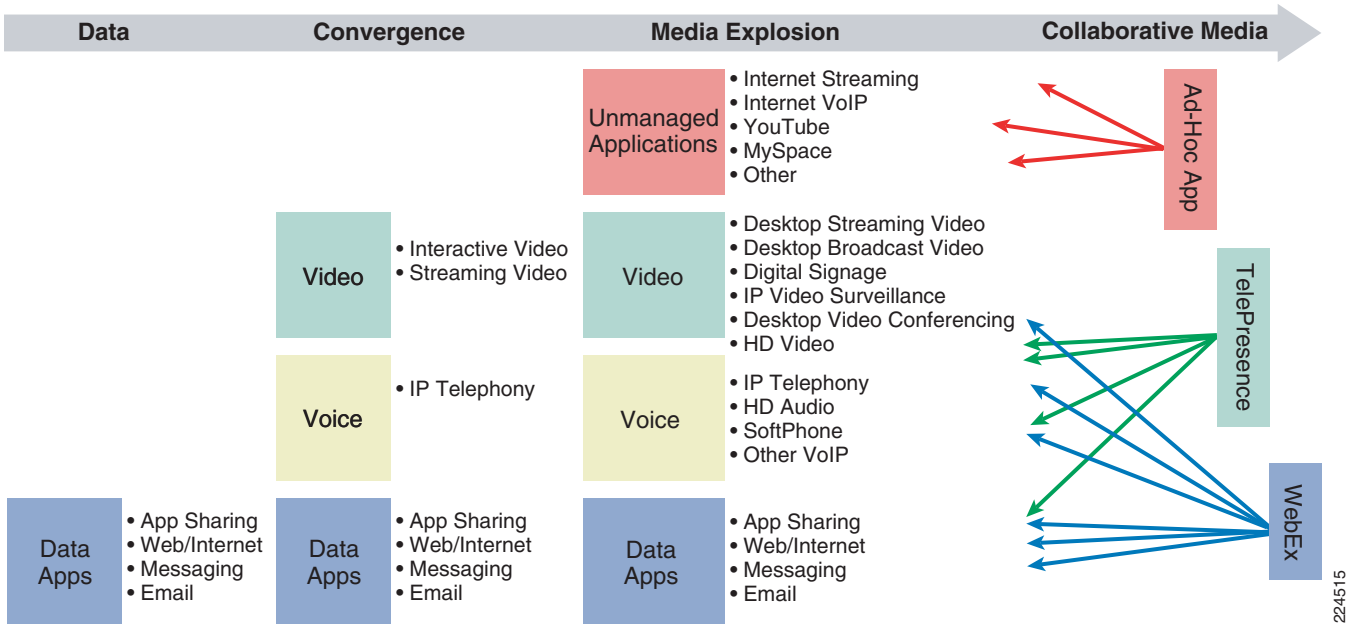
The Convergence Within Media Applications

Much like the integration of rich text and graphics into documentation, audio and video media continue to be integrated into many forms of communication. Sharing of information with E-mailed slide sets will gradually be replaced with E-mailed video clips. The audio conference bridge will be supplanted with the video-enabled conference bridge. Collaboration tools designed to link together distributed employees will increasingly integrate desktop video to bring teams closer together.

Cisco WebEx is a prime example of such integration, providing text, audio, instant messaging, application sharing, and desktop video conferencing easily to all meeting participates, regardless of their location. Instead of a cumbersome setup of a video conference call, applications such as CUPC and WebEx greatly simplify the process and video capability is added to the conference just as easily as any other type of media, such as audio.

The complexity that application presents to the network administrator relates to application classification: as media applications include voice, video, and data sub-components, the question of how to mark and provision a given media application becomes more difficult and blurry, as illustrated in Figure 4-3.

Figure 4-3 Media Application Convergence—Voice, Video, and Data Within an Application



For example, since Cisco WebEx has voice, video, and data sub-components, how should it be classified? As a voice application? As a video application? As a data application? Or is an altogether new application-class model needed to accommodate multimedia applications?

The Globalization of the Workforce

In the past, most companies focused on acquiring and retaining skilled and talented individuals in a single or few geographic locations. More recently, this focus has shifted to finding technology solutions to enable a geographically-distributed workforce to collaborate together as a team. This new approach enables companies to more flexibly harness talent "where it lives."

Future productivity gains will be achieved by creating collaborative teams that span corporate boundaries, national boundaries, and geographies. Employees will collaborate with partners, research and educational institutions, and customers to create a new level of collective knowledge.

To do so, real-time multimedia collaboration applications are absolutely critical to the success of these virtual teams. Video offers a unique medium which streamlines the effectiveness of communications between members of such teams. For this reason, real-time interactive video will become increasingly prevalent, as will media integrated with corporate communications systems.

The Pressures to be Green

For many reasons, companies are seeking to reduce employee travel. Travel creates bottom line expenses, as well as significant productivity impacts while employees are in-transit and away from their usual working environments. Many solutions have emerged to assist with productivity while traveling, including wireless LAN hotspots, remote access VPNs, and softphones, all designed to keep the employee connected while traveling.

More recently companies are under increasing pressures to demonstrate environmental responsibility, often referred to as being "green." On the surface, such initiatives may seem like a pop-culture trend that lacks tangible corporate returns. However, it is entirely possible to pursue green initiatives while simultaneously increasing productivity and lowering expenses.

Media applications, such as Cisco TelePresence, offer real solutions to remote collaboration challenges and have demonstrable savings as well. For example, during the first year of deployment, Cisco measured its usage of TelePresence in direct comparison to the employee travel that would otherwise have taken place. Cisco discovered that over 80,000 hours of meetings were held by TelePresence instead of physical travel, avoiding $100 million of travel expenses, as well as over 30,000 tons of carbon emissions, the equivalent of removing over 10,000 vehicles off the roads for a period of one year.

Being green does not have to be a "tax;" but rather can improve productivity and reduce corporate expenses, offering many dimensions of return on investment, while at the same time sending significant messages to the global community of environmental responsibility.

Thus, having reviewed several key business drivers for evolving QoS designs, relevant industry guidance and best practices are discussed next.

New Industry Guidance and Best Practices

A second set of drivers behind QoS design evolution are advances in industry standards and guidance. Cisco has long advocated following industry standards and recommendations—whenever possible—when deploying QoS, as this simplifies QoS designs, extends QoS policies beyond an administrative domain, and improves QoS between administrative domains.

To the first point of simplifying QoS, there are 64 discrete Differentiated Services Code Point (DSCP) values to which IP packets can be marked. If every administrator were left to their own devices to arbitrarily pick-and-choose DSCP markings for applications, there would be a wide and disparate set of

marking schemes that would likely vary from enterprise to enterprise, perhaps even within an enterprise (such as department to department). However if industry standard marking values are used, then marking schemes become considerably simplified and consistent.

To the second point of extending QoS policies beyond an administrative domain, if an enterprise administrator wishes a specific type of Per-Hop Behavior (PHB)—which is the manner in which a packet marked to a given DSCP value is to be treated at each network node—they mark the packet according to the industry recommended marking value that corresponds to the desired PHB. Then, as packets are handed off to other administrative domains, such as service provider networks or partner networks, these packets continue to receive the desired PHB (provided that the SP or partner network is also following the same industry standards). Therefore, the PHB treatment is extended beyond the original administrative domain and thus the overall quality of service applied to the packet end-to-end-is improved.

To the third point of improving QoS between administrative domains, as networks pass packets to adjacent administrative domains, sometimes their QoS policies differ. Nonetheless, the differences are likely to be minor, as compared to the scenario in which every administrator handled packets in an arbitrary, locally-defined fashion. Thus, the mapping of QoS policies is much easier to handle between domains, as these ultimately use many—if not most—of the same industry-defined PHBs.

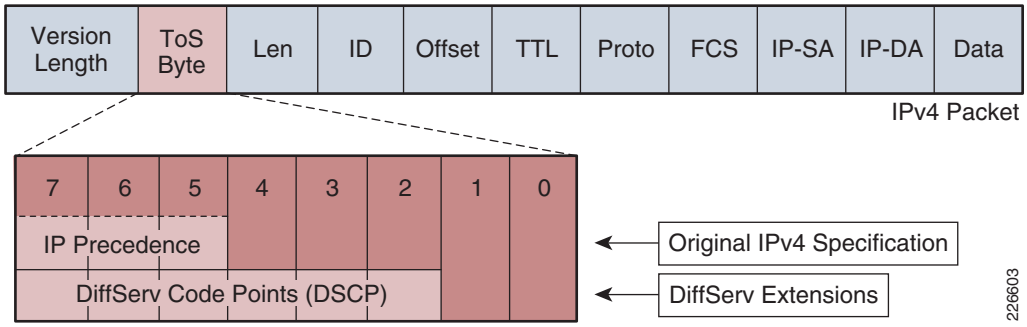
However, there may be specific constraints, either financial, technical, or otherwise, that may preclude following industry standards 100% of the time. In such cases, administrators need to make careful decisions as to when and how to deviate from these standards and recommendations to best meet their specific objectives and constraints and to allow them maximum flexibility and consistency in the end-to-end scenarios described above.

Therefore, in line with the principle of following industry standards and recommendations whenever possible, it would be beneficial to briefly review some of the standards and recommendations most relevant to QoS design.

RFC 2474 Class Selector Code Points

The IETF RFC 2474 standard defines the use of 6 bits in the IPv4 and IPv6 Type of Service (ToS) byte, termed Differentiated Services Code Points (DSCP). Additionally, this standard introduces Class Selector codepoints to provide backwards compatibility for legacy (RFC 791) IP Precedence bits, as shown in Figure 4-4.

Figure 4-4 The IP ToS Byte—IP Precedence Bits and DiffServ Extensions



Class Selectors, defined in RFC 2474, are not Per-Hop Behaviors per se, but rather were defined to provide backwards compatibility to IP Precedence. Each Class Selector corresponds to a given IP Precedence value, with its three least-significant bits set to 0. For example, IP Precedence 1 is referred to as Class Selector 1 (or DSCP 8), IP Precedence 2 is referred to as Class Selector 2 (or DSCP 16), and so on. Table 4-1 shows the full table of IP Precedence to Class Selector mappings.

Table 4-1 IP Precedence to Class Selector/DSCP Mappings

IP Precedence Value	IP Precedence Name	IPP Binary Equivalent	Class Selector	CS Binary Equivalent	DSCP Value (Decimal)
0	Normal	000	CS0[1]/DF	000 000	0
1	Priority	001	CS1	001 000	8
2	Immediate	010	CS2	010 000	16
3	Flash	011	CS3	011 000	24
4	Flash-Override	100	CS4	100 000	32
5	Critical	101	CS5	101 000	40
6	Internetwork Control	110	CS6	110 000	48
7	Network Control	111	CS7	111 000	56

1. Class Selector 0 is a special case, as it represents the default marking value (defined in RFC 2474-Section 4.1); as such, it is not typically called Class Selector 0, but rather Default Forwarding or DF.

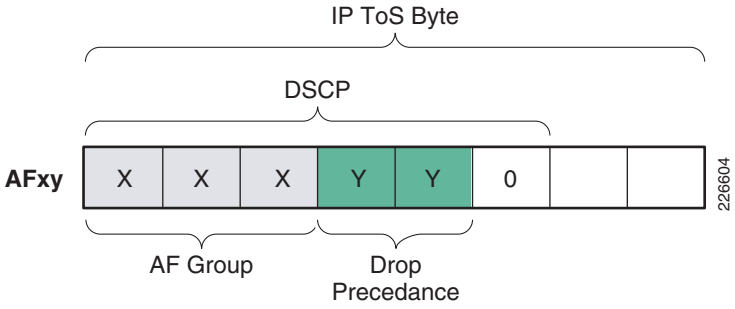
RFC 2597 Assured Forwarding Per-Hop Behavior Group

RFC 2597 defines four Assured Forwarding groups, denoted by the letters "AF" followed by two digits:

- The first digit denotes the AF class number and can range from 1 through 4 (these values correspond to the three most-significant bits of the codepoint or the IPP value that the codepoint falls under). Incidentally, the AF class number does not in itself represent hierarchy (that is, AF class 4 does not necessarily get any preferential treatment over AF class 1).
- The second digit refers to the level of drop precedence within each AF class and can range from 1 (lowest drop precedence) through 3 (highest drop precedence).

Figure 4-5 shows the Assured Forwarding PHB marking scheme.

Figure 4-5 Assured Forwarding PHB Marking Scheme



The three levels of drop precedence are analogous to the three states of a traffic light:

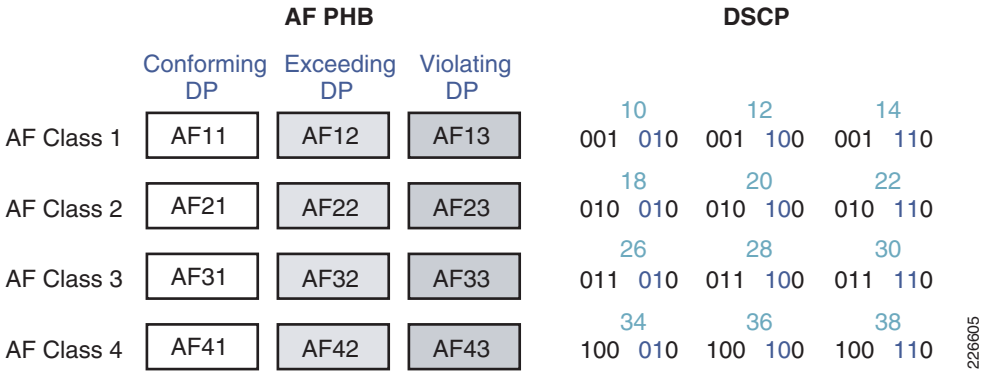
- Drop precedence 1, also known as the "conforming" state, is comparable to a green traffic light.
- Drop precedence 2, also known as the "exceeding" state, is comparable to a yellow traffic light (where a moderate allowance in excess of the conforming rate is allowed to prevent erratic traffic patterns).
- Drop precedence 3, also known as the "violating" state, is comparable to a red traffic light.

Packets within an AF class are always initially marked to drop precedence of 1 and can only be remarked to drop precedence 2 or 3 by a policer, which meters traffic rates and determines if the traffic is exceeding or violating a given traffic contract.

Then, for example, during periods of congestion on an RFC 2597-compliant node, packets remarked AF33 (representing the highest drop precedence for AF class 3) would be dropped more often than packets remarked AF32; in turn, packets remarked AF32 would be dropped more often than packets marked AF31.

The full set of AF PHBs are detailed in Figure 4-6.

Figure 4-6 Assured Forwarding PHBs with Decimal and Binary Equivalents



RFC 3246 An Expedited Forwarding Per-Hop Behavior

The Expedited Forwarding PHB is defined in RFC 3246. In short, the definition describes a strict-priority treatment for packets that have been marked to a DSCP value of 46 (101110), which is also termed Expedited Forwarding (or EF). Any packet marked 46/EF that encounters congestion at a given network node is to be moved to the front-of-the-line and serviced in a strict priority manner. It doesn't matter how such behavior is implemented—whether in hardware or software—as long as the behavior is met for the given platform at the network node.

Note Incidentally, the RFC 3246 does not specify which application is to receive such treatment; this is open to the network administrator to decide, although the industry norm over the last decade has been to use the EF PHB for VoIP.

The EF PHB provides an excellent case-point of the value of standardized PHBs. For example, if a network administrator decides to mark his VoIP traffic to EF and service it with strict priority over his networks, he can extend his policies to protect his voice traffic even over networks that he does not have direct administrative control. He can do this by partnering with service providers and/or extranet partners who follow the same standard PHB and who thus continue to service his (EF marked) voice traffic with strict priority over their networks.

RFC 3662 A Lower Effort Per-Domain Behavior for Differentiated Services

While most of the PHBs discussed so far represent manners in which traffic may be treated preferentially, there are cases it may be desired to treat traffic deferentially. For example, certain types of non-business traffic, such as gaming, video-downloads, peer-to-peer media sharing, and so on might dominate network links if left unabated.

To address such needs, a Lower Effort Per-Domain Behavior is described in RFC 3662 to provide a less than Best Effort service to undesired traffic. Two things should be noted about RFC 3662 from the start:

- RFC 3662 is in the "informational" category of RFCs (not the standards track) and as such is not necessary to implemented in order to be DiffServ standard-compliant.
- A Per-Domain Behavior (PDB) has a different and larger scope than a Per-Hop Behavior (PHB). A PDB does not require that undesired traffic be treated within a "less than Best Effort service" at necessarily every network node (which it would if this behavior were defined as a Per-Hop Behavior); rather, as long as one (or more) nodes within the administrative domain provide a "less than best effort service" to this undesired traffic class, the Per-Domain Behavior requirement has been met.

The reason a PDB is sufficient to provision this behavior, as opposed to requiring a PHB, is that the level of service is deferential, not preferential. To expand, when dealing with preferential QoS policies, sometimes it is said that "a chain of QoS policies is only as strong as the weakest link." For example, if provisioning an EF PHB for voice throughout a network and only one node in the path does not have EF properly provisioned on it, then the overall quality of voice is (potentially) ruined. On the other hand, if the objective is to provide a deferential level of service, all one needs is a single weak link in the path in order to lower the overall quality of service for a given class. Thus, if only a single weak link is required per administrative domain, then a Per-Domain Behavior-rather than a Per-Hop Behavior-better suits the requirement.

The marking value recommended in RFC 3662 for less than best effort service (sometimes referred to as a Scavenger service) is Class Selector 1 (DSCP 8). This marking value is typically assigned and constrained to a minimally-provisioned queue, such that it will be dropped the most aggressively under network congestion scenarios.

Cisco's QoS Baseline

While the IETF DiffServ RFCs (discussed thus far) provided a consistent set of per-hop behaviors for applications marked to specific DSCP values, they never specified which application should be marked to which DiffServ Codepoint value. Therefore, considerable industry disparity existed in application-to-DSCP associations, which led Cisco to put forward a standards-based application marking recommendation in their strategic architectural QoS Baseline document (in 2002). Eleven different application classes were examined and extensively profiled and then matched to their optimal RFC-defined PHBs. The application-specific marking recommendations from Cisco's QoS Baseline of 2002 are summarized in Figure 4-7.

Figure 4-7 Cisco's QoS Baseline Marking Recommendations

Application	L3 Classification		IETF
	PHB	DSCP	RFC
Routing	CS6	48	RFC 2474
Voice	EF	46	RFC 3246
Interactive Video	AF41	34	RFC 2597
Streaming Video	CS4	32	RFC 2474
Mission-Critical Data	AF31	26	RFC 2597
Call Signaling	CS3	24	RFC 2474
Transactional Data	AF21	18	RFC 2597
Network Management	CS2	16	RFC 2474
Bulk Data	AF11	10	RFC 2597
Best Effort	0	0	RFC 2474
Scavenger	CS1	8	RFC 2474

220199

Note The previous Cisco Enterprise QoS SRND (version 3.3 from 2003) was based on Cisco's QoS Baseline; however, as will be discussed, newer RFCs have since been published that improve and expand on the Cisco QoS Baseline.

RFC 4594 Configuration Guidelines for DiffServ Classes

More than four years after Cisco put forward its QoS Baseline document, RFC 4594 was formally accepted as an informational RFC (in August 2006).

Before getting into the specifics of RFC 4594, it is important to comment on the difference between the IETF RFC categories of informational and standard. An informational RFC is an industry recommended best practice, while a standard RFC is an industry requirement. Therefore RFC 4594 is a set of formal DiffServ QoS configuration best practices, not a requisite standard.

RFC 4594 puts forward twelve application classes and matches these to RFC-defined PHBs. These application classes and recommended PHBs are summarized in Figure 4-8.

Figure 4-8 RFC 4594 Marking Recommendations

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

220200

It is fairly obvious that there are more than a few similarities between Cisco's QoS Baseline and RFC 4594, as there should be, since RFC 4594 is essentially an industry-accepted evolution of Cisco's QoS Baseline. However, there are some differences that merit attention.

The first set of differences is minor, as they involve mainly nomenclature. Some of the application classes from the QoS Baseline have had their names changed in RFC 4594. These changes in nomenclature are summarized in Table 4-2.

Table 4-2 Nomenclature Changes from Cisco QoS Baseline to RFC 4594

Cisco QoS Baseline Class Names	RFC 4594 Class Names
Routing	Network Control
Voice	VoIP Telephony
Interactive Video	Multimedia Conferencing
Streaming Video	Multimedia Streaming
Transactional Data	Low-Latency Data
Network Management	Operations/Administration/Management (OAM)
Bulk Data	High-Throughput Data
Scavenger	Low-Priority Data

The remaining changes are more significant. These include one application class deletion, two marking changes, and two new application class additions:

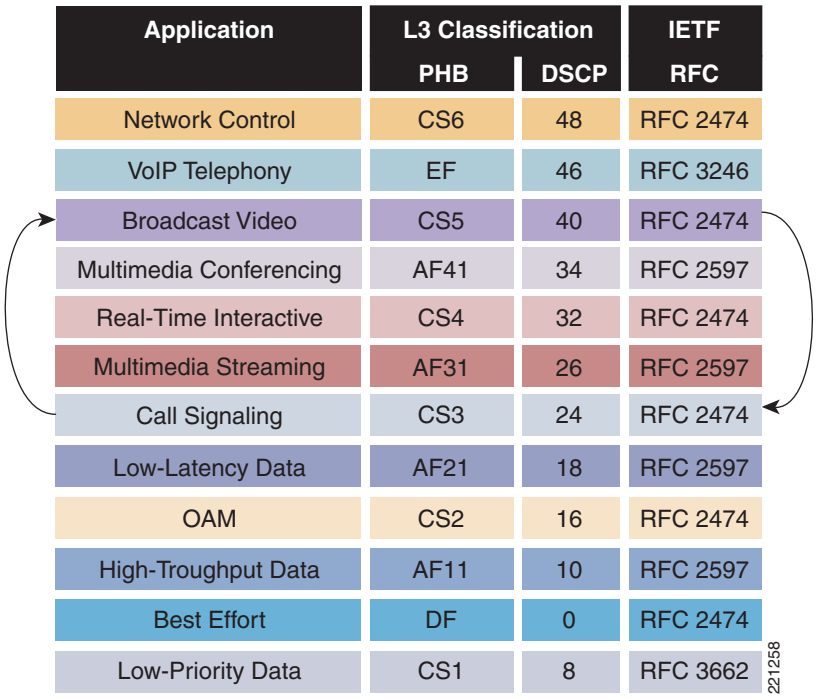
- The QoS Baseline Locally-Defined Mission-Critical Data class has been deleted from RFC 4594.
- The QoS Baseline marking recommendation of CS4 for Streaming Video has been changed in RFC 4594 to mark Multimedia Streaming to AF31.

- The QoS Baseline marking recommendation of CS3 for Call Signaling has been changed in RFC 4594 to mark Call Signaling to CS5.
- A new application class has been added to RFC 4594, Real-Time Interactive. This addition allows for a service differentiation between elastic conferencing applications (which would be assigned to the Multimedia Conferencing class) and inelastic conferencing applications (which would include high-definition applications, like Cisco TelePresence, in the Realtime Interactive class). Elasticity refers to the applications ability to function despite experiencing minor packet loss. Multimedia Conferencing uses the AF4 class and is subject to markdown (and potential dropping) policies, while the Realtime Interactive class uses CS4 and is not subject neither to markdown nor dropping policies.
- A second new application class has been added to RFC 4594, Broadcast video. This addition allows for a service differentiation between elastic and inelastic streaming media applications. Multimedia Streaming uses the AF3 class and is subject to markdown (and potential dropping) policies, while Broadcast Video uses the CS3 class and is subject neither to markdown nor dropping policies.

The most significant of the differences between Cisco's QoS Baseline and RFC 4594 is the RFC 4594 recommendation to mark Call Signaling to CS5. Cisco has completed a lengthy and expensive marking migration for Call Signaling from AF31 to CS3 (as per the original QoS Baseline of 2002) and, as such, there are no plans to embark on another marking migration in the near future. It is important to remember that RFC 4594 is an informational RFC (in other words, an industry best-practice) and not a standard. Therefore, lacking a compelling business case at the time of writing, Cisco plans to continue marking Call Signaling as CS3 until future business requirements arise that necessitate another marking migration.

Therefore, for the remainder of this document, RFC 4594 marking values are used throughout, with the one exception of swapping Call-Signaling marking (to CS3) and Broadcast Video (to CS5). These marking values are summarized in Figure 4-9.

Figure 4-9 Cisco-Modified RFC 4594-based Marking Values (Call-Signaling is Swapped with Broadcast Video)



A final note regarding standards and RFCs is that there are other RFCs relating to DiffServ design that are currently in draft status (as of the time of writing). One such example is RFC 5127, "Aggregation of Diffserv Service Classes." As such drafts are finalized, these will correspondingly impact respective areas of QoS design.

Having reviewed various relevant industry guidance and best practices relating to QoS evolution, a final driver—namely advances in QoS technologies—is briefly introduced.

New Platforms and Technologies

As network hardware and software technologies evolve, so do their QoS capabilities and features. New switches and linecards boast advanced classification engines or queuing structures, new routers support sophisticated QoS tools that scale with greater efficiency, and new IOS software features present entirely new QoS options to solve complex scenarios. Therefore, a third set of drivers behind QoS design evolution are the advances in QoS technologies, which are discussed in detail in their respective Place-in-the-Network (PIN) QoS design chapters.

As can be noted from the discussion to this point, all of the drivers behind QoS design evolution are in a constant state of evolution themselves—business drivers will continue to expand and change, as will relevant industry standards and guidance, and so too will platforms and technologies. Therefore, while the strategic and detailed design recommendations presented in this document are as forward-looking as possible, these will no doubt continue to evolve over time.

Before discussing current strategic QoS design recommendations, it may be beneficial to set a base context by first overviewing Cisco's QoS toolset.

Cisco QoS Toolset

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

- Admission control tools
- Classification and marking tools
- Policing and markdown tools
- Scheduling tools
- Link-efficiency tools
- Hierarchical QoS
- AutoQoS
- QoS management

Classification and Marking Tools

Classification tools serve to identify traffic flows so that specific QoS actions may be applied to the desired flows. Often the terms classification and marking are used interchangeably (yet incorrectly so); therefore, it is important to understand the distinction between classification and marking operations:

- Classification refers to the inspection of one or more fields in a packet (the term packet is being used loosely here, to include all Layer 2 to Layer 7 fields, not just Layer 3 fields) to identify the type of traffic that the packet is carrying. Once identified, the traffic is directed to the applicable

policy-enforcement mechanism for that traffic type, where it receives predefined treatment (either preferential or deferential). Such treatment can include marking/remarking, queuing, policing, shaping, or any combination of these (and other) actions.

- Marking, on the other hand, refers to changing a field within the packet to preserve the classification decision that was reached. Once a packet has been marked, a "trust-boundary" is established on which other QoS tools later depend. Marking is only necessary at the trust boundaries of the network and (as with all other QoS policy actions) cannot be performed without classification. By marking traffic at the trust boundary edge, subsequent nodes do not have to perform the same in-depth classification and analysis to determine how to treat the packet.

Cisco IOS software performs classification based on the logic defined within the class map structure within the Modular QoS Command Line Interface (MQC) syntax. MQC class maps can perform classification based on the following types of parameters:

- Layer 1 parameters—Physical interface, sub-interface, PVC, or port
- Layer 2 parameters—MAC address, 802.1Q/p Class of Service (CoS) bits, MPLS Experimental (EXP) bits
- Layer 3 parameters—Differentiated Services Code Points (DSCP), IP Precedence (IPP), IP Explicit Congestion Notification (IP ECN), source/destination IP address
- Layer 4 parameters—TCP or UDP ports
- Layer 7 parameters—Application signatures and URLs in packet headers or payload via Network Based Application Recognition (NBAR)

NBAR is the most sophisticated classifier in the IOS tool suite. NBAR can recognize packets on a complex combination of fields and attributes. NBAR deep-packet classification engine examines the data payload of stateless protocols and identifies application-layer protocols by matching them against a Protocol Description Language Module (PDLM), which is essentially an application signature. NBAR is dependent on Cisco Express Forwarding (CEF) and performs deep-packet classification only on the first packet of a flow. The rest of the packets belonging to the flow are then CEF-switched. However, it is important to recognize that NBAR is merely a classifier, nothing more. NBAR can identify flows by performing deep-packet inspection, but it is up to the MQC policy-map to define what action should be taken on these NBAR-identified flows.

Marking tools change fields within the packet, either at Layer 2 or at Layer 3, such that in-depth classification does not have to be performed at each network QoS decision point. The primary tool within MQC for marking is Class-Based Marking (though policers-sometimes called markers-may also be used, as is discussed shortly). Class-Based Marking can be used to set the CoS fields within an 802.1Q/p tag (as shown in Figure 4-10), the Experimental bits within a MPLS label (as shown in Figure 4-11), the Differentiated Services Code Points (DSCPs) within an IPv4 or IPv6 header (as shown in Figure 4-12), the IP ECN Bits (also shown in Figure 4-12), as well as other packet fields. Class-Based Marking, like NBAR, is CEF-dependant.

Figure 4-10 Figure 1-10 802.1Q/p CoS Bits

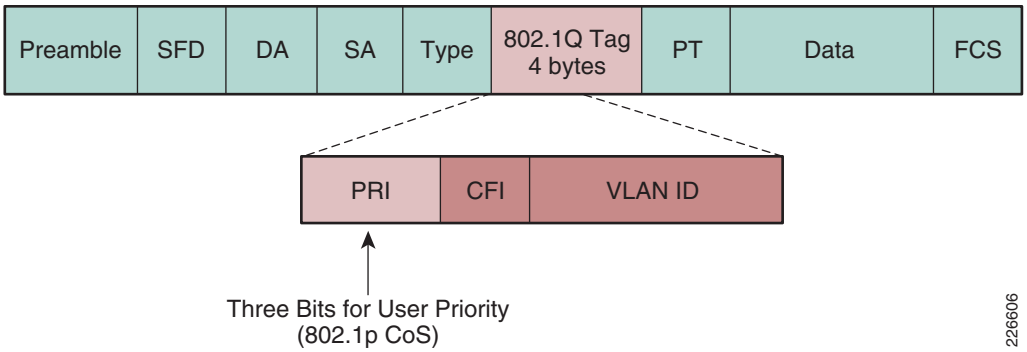


Figure 4-11 Figure 1-11 MPLS EXP Bits

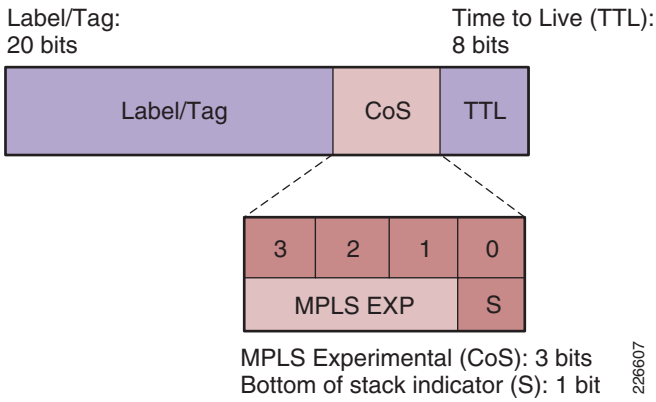
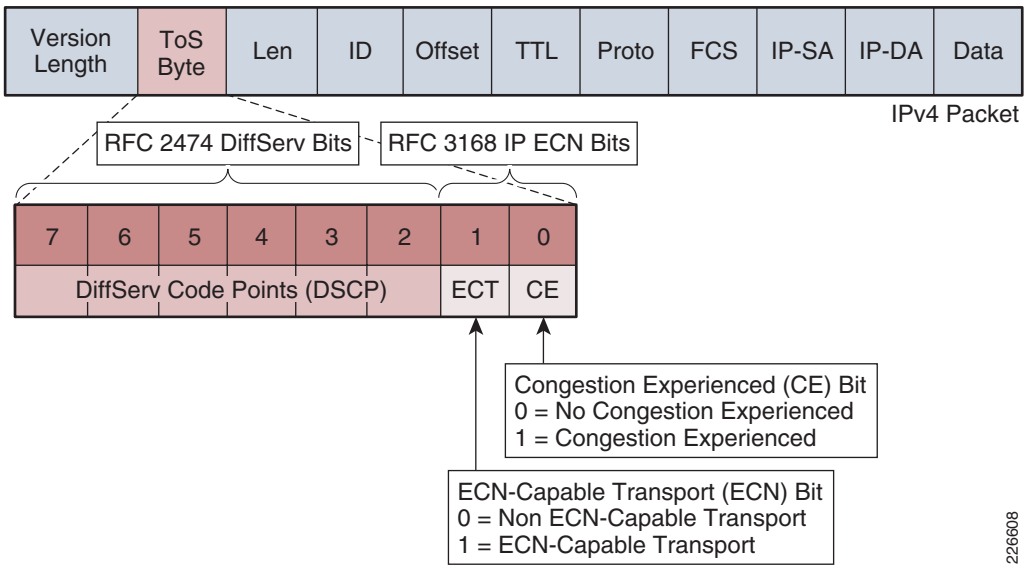


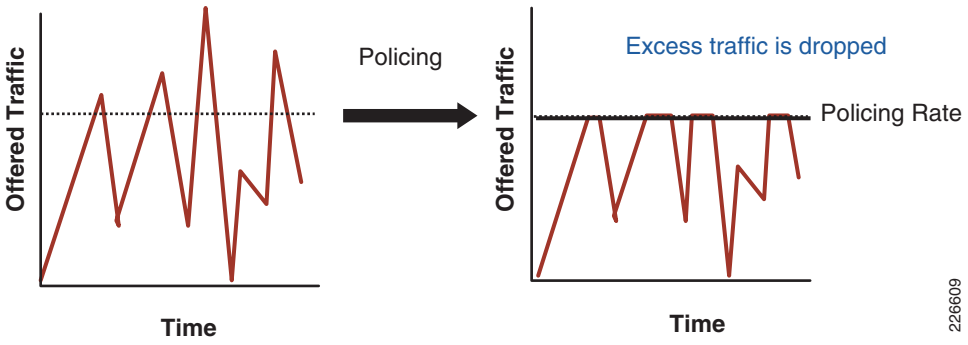
Figure 4-12 Figure 1-12 IP ToS Bits: DSCP and IP ECN



Policing and Markdown Tools

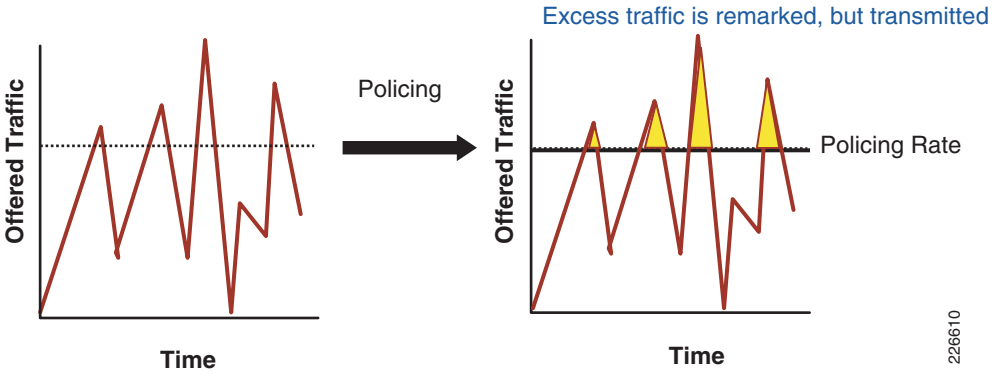
Policers are used to monitor traffic flows and to identify and respond to traffic violations. Policers achieve these objectives by performing ongoing, instantaneous checks for traffic violations and taking immediate prescribed actions when such violations occur. For example, a policer can determine if the offered load is in excess of the defined traffic rate and then drop the out-of-contract traffic, as illustrated in Figure 4-13.

Figure 4-13 A Policing Action



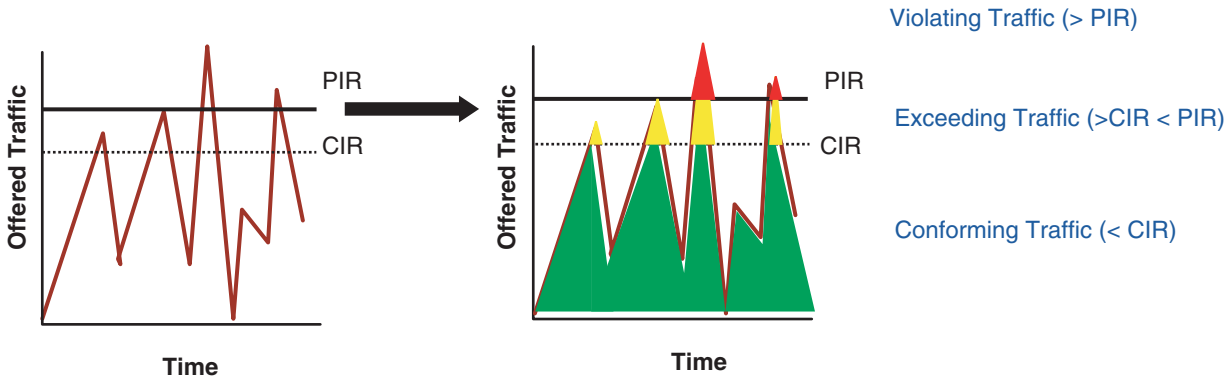
Alternatively, policers may be used to remark excess traffic instead of dropping it. In such a role, the policer is called a marker. Figure 4-14 illustrates a policer functioning as a marker.

Figure 4-14 A Policer as a Marker



The rate at which the policer is configured to either drop or remark traffic is called the Committed Information Rate (CIR). However, policers may police to multiple rates, such as the dual rate policer defined in RFC 2698. With such a policer, the CIR is the principle rate to which traffic is policed, but an upper limit, called the Peak Information Rate (PIR), is also set. The action of a dual-rate policer is analogous to a traffic light, with three conditional states—green light, yellow light, and red light. Traffic equal to or below the CIR (a green light condition) is considered to conform to the rate. An allowance for moderate amounts of traffic above this principal rate is permitted (a yellow light condition) and such traffic is considered to exceed the rate. However, a clearly-defined upper-limit of tolerance (the PIR) is also set (a red light condition), beyond which traffic is considered to violate the rate. As such, a dual-rate RFC 2698 policer performs the traffic conditioning for RFC 2597 Assured Forwarding PHBs, as previously discussed. The actions of such a dual-rate policer (functioning as a three-color marker) are illustrated in Figure 4-15.

Figure 4-15 A Dual-Rate Policer as a Three-Color Marker



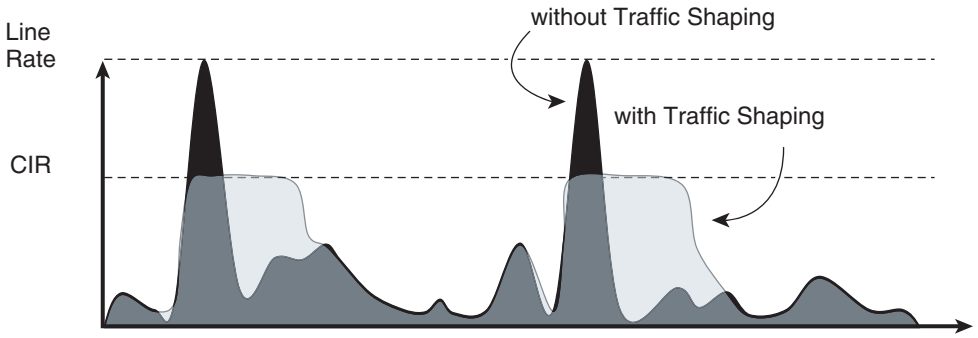
A RFC 2698 “Two Rate Three Color Marker” can:
- mark *conforming* traffic to one value (such as AF31)
- remark *exceeding* traffic to another value (such as AF32)
- remark *violating* traffic to yet another value (such as AF33)

226611

Shaping Tools

Shapers operate in a manner similar to policers, in that they meter traffic rates. However, the principle difference between a policer and a shaper is that where a policer remarks or drops traffic as a policy action, a shaper merely delays traffic. Figure 4-16 illustrates generic traffic shaping.

Figure 4-16 Traffic Shaping



Traffic shaping limits the transmit rate of traffic
to a value (CIR) lower than the interface’s line rate
by temporarily buffering packets exceeding the CIR

226612

Shapers are particularly useful when traffic must conform to a specific rate of traffic in order to meet a service level agreement (SLA) or to guarantee that traffic offered to a service provider is within a contracted rate. Traditionally, shapers have been associated with Non-Broadcast Multiple-Access (NBMA) Layer 2 WAN topologies, like ATM and Frame-Relay, where potential speed-mismatches exist. However, shapers are becoming increasingly necessary on Layer 3 WAN access circuits, such as Ethernet-based handoffs, in order to conform to sub-line access-rates.

Queuing and Dropping Tools

Normally, over uncongested interfaces, packets are transmitted in order on a First-In-First-Out (FIFO) basis. However, if packets arrive at an interface faster than they can be transmitted out the interface, then excess packets may be buffered. When packets are buffered, they may be reordered prior to transmission according to administratively-defined algorithms, which are generally referred to as queuing policies. It is important to recognize that queuing policies are engaged only when the interface is experiencing congestion and are deactivated shortly after the interface congestion clears.

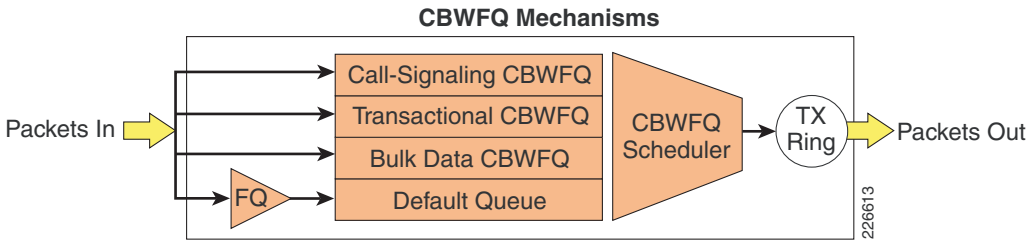
Queuing may be performed in software or in hardware. Within Cisco IOS Software there are two main queuing algorithms available, Class-Based Weighted-Fair Queuing (CBWFQ) and Low-Latency Queuing (LLQ). Within Cisco Catalyst hardware, queuing algorithms fall under a 1PxQyT model, which are overviewed in the following sections.

CBWFQ

Regardless of what queuing policy is applied to an interface within Cisco IOS, there is always an underlying queuing mechanism in place called the Tx-Ring, which is a final (FIFO) output buffer. The Tx-Ring serves the purpose of always having packets ready to be placed onto the wire so that link utilization can be driven to 100%. The Tx-Ring also serves to indicate congestion to the IOS software; specifically, when the Tx-Ring fills to capacity, then the interface is known to be congested and a signal is sent to engage any LLQ/CBWFQ policies that have been configured on the interface.

Class-Based Weighted-Fair Queuing (CBWFQ) is a queuing algorithm that combines the ability to guarantee bandwidth with the ability to dynamically ensure fairness to other flows within a class of traffic. Each queue is serviced in a weighted-round-robin (WRR) fashion based on the bandwidth assigned to each class. The operation of CBWFQ is illustrated in Figure 4-17.

Figure 4-17 CBWFQ Operation



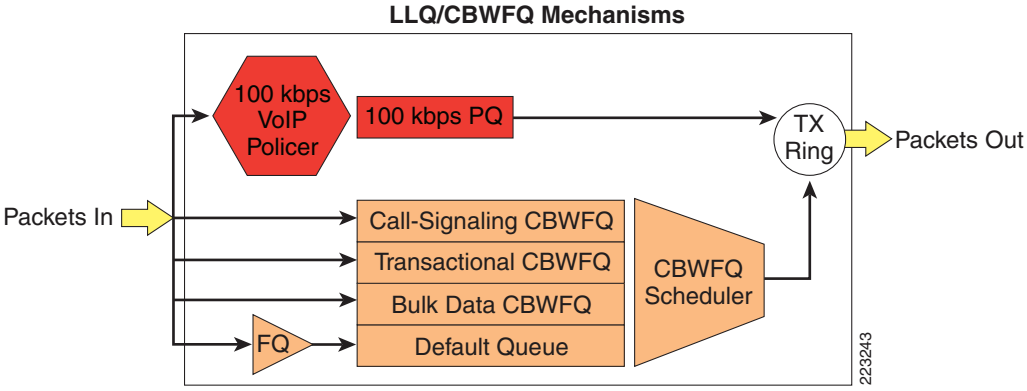
In Figure 4-17, a router interface has been configured with a 4-class CBWFQ policy, with an explicit CBWFQ defined for Network Control, Transactional Data, and Bulk Data respectively, as well as the default CBWFQ queue, which has a Fair-Queuing (FQ) pre-sorter assigned to it.

Note CBWFQ is a bit of a misnomer because the pre-sorter that may be applied to certain CBWFQs, such as class-default, is not actually a Weighted-Fair Queuing (WFQ) pre-sorter, but rather a Fair-Queuing (FQ) pre-sorter. As such, it ignores any IP Precedence values when calculating bandwidth allocations traffic flows. To be more technically precise, this queuing algorithm would be more accurately named Class-Based Fair-Queuing or CBFQ.

LLQ

Low-Latency Queuing (LLQ) is essentially CBWFQ combined with a strict priority queue. In fact, the original name for the LLQ scheduling algorithm was PQ-CBWFQ. While this name was technically more descriptive, it was obviously clumsy from a marketing perspective and hence the algorithm was renamed LLQ. LLQ operation is illustrated in Figure 4-18.

Figure 4-18 LLQ/CBWFQ Operation

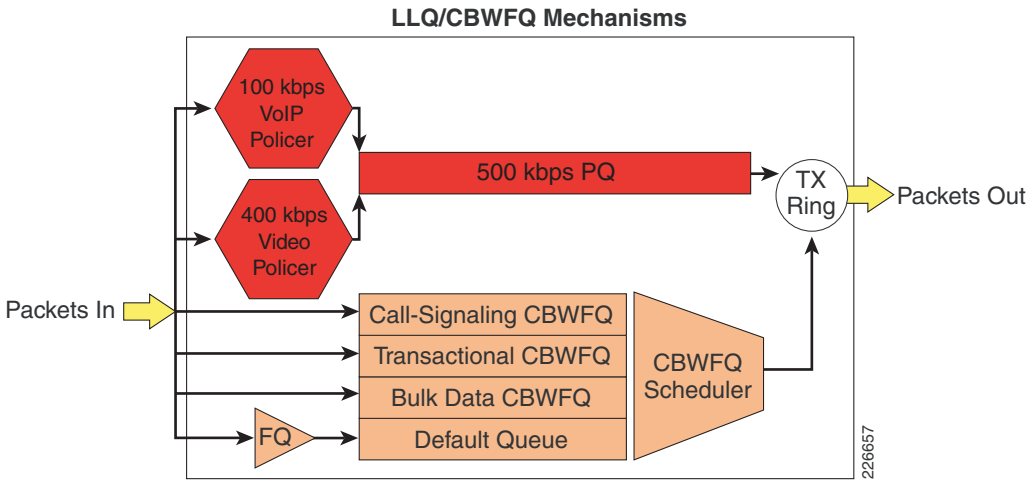


In Figure 4-18, a router interface has been configured with a 5-class LLQ/CBWFQ policy, with voice assigned to a 100 kbps LLQ, three explicit CBWFQs are defined for Call-Signaling, Transactional Data, and Bulk Data respectively, as well as a default queue that has a Fair-Queuing pre-sorter assigned to it. However, an underlying mechanism that doesn't appear within the IOS configuration, but is shown in Figure 4-18, is an implicit policer attached to the LLQ.

The threat posed by any strict priority-scheduling algorithm is that it could completely starve lower priority traffic. To prevent this, the LLQ mechanism has a built-in policer. This policer (like the queuing algorithm itself) engages only when the LLQ-enabled interface is experiencing congestion. Therefore, it is important to provision the priority classes properly. In this example, if more than 100 kbps of voice traffic was offered to the interface, and the interface was congested, the excess voice traffic would be discarded by the implicit policer. However, traffic that is admitted by the policer gains access to the strict priority queue and is handed off to the Tx-Ring ahead of all other CBWFQ traffic.

Not only does the implicit policer for the LLQ protect CBWFQs from bandwidth-starvation, but it also allows for sharing of the LLQ. TDM of the LLQ allows for the configuration and servicing of multiple LLQs, while abstracting the fact that there is only a single LLQ "under-the-hood," so to speak. For example, if both voice and video applications required realtime service, these could be provisioned to two separate LLQs, which would not only protect voice and video from data, but also protect voice and video from interfering with each other, as illustrated in Figure 4-19.

Figure 4-19 Dual-LLQ/CBWFQ Operation



In Figure 4-19, a router interface has been configured with a 6-class LLQ/CBWFQ policy, with voice assigned to a 100 kbps LLQ, video assigned to a "second" 400 kbps LLQ, three explicit CBWFQs are defined for Call-Signaling, Transactional Data, and Bulk Data respectively, as well as a default queue that has a Fair-Queuing pre-sorter assigned to it.

Within such a dual-LLQ policy, two separate implicit policers have been provisioned, one each for the voice class (to 100 kbps) and another for the video class (to 400 kbps), yet there remains only a single strict-priority queue, which is provisioned to the sum of all LLQ classes, in this case to 500 kbps (100 kbps + 400 kbps). Traffic offered to either LLQ class is serviced on a first-come, first-serve basis until the implicit policer for each specific class has been invoked. For example, if the video class attempts to burst beyond its 400 kbps rate then it is dropped. In this manner, both voice and video are serviced with strict-priority, but do not starve data flows, nor do they interfere with each other.

1PxQyT

In order to scale QoS functionality to campus speeds (like GigabitEthernet or Ten GigabitEthernet), Catalyst switches must perform QoS operations within hardware. For the most part, classification, marking, and policing policies (and syntax) are consistent in both Cisco IOS Software and Catalyst hardware; however, queuing (and dropping) policies are significantly different when implemented in hardware. Hardware queuing across Catalyst switches is implemented in a model that can be expressed as 1PxQyT, where:

- 1P represents the support of a strict-priority hardware queue (which is usually disabled by default).
- *x*Q represents *x* number of non-priority hardware queues (including the default, Best-Effort queue).
- *y*T represents *y* number of drop-thresholds per non-priority hardware queue.

For example, consider a Catalyst 6500 48-port 10/100/1000 RJ-45 Module, the WS-X6748-GE-TX, which has a 1P3Q8T egress queuing structure, meaning that it has:

- One strict priority hardware queue
- Three additional non-priority hardware queues, each with:
  - Eight configurable Weighted Random Early Detect (WRED) drop thresholds per queue

Traffic assigned to the strict-priority hardware queue is treated with an Expedited Forwarding Per-Hop Behavior (EF PHB). That being said, it bears noting that on some platforms there is no explicit limit on the amount of traffic that may be assigned to the PQ and as such, the potential to starve non-priority

queues exists. However, this potential for starvation may be effectively addressed by explicitly configuring input policers that limit—on a per-port basis—the amount of traffic that may be assigned to the priority queue (PQ). Incidentally, this is the recommended approach defined in RFC 3246 (Section 3).

Traffic assigned to a non-priority queue is provided with bandwidth guarantees, subject to the PQ being either fully-serviced or bounded with input policers.

WRED

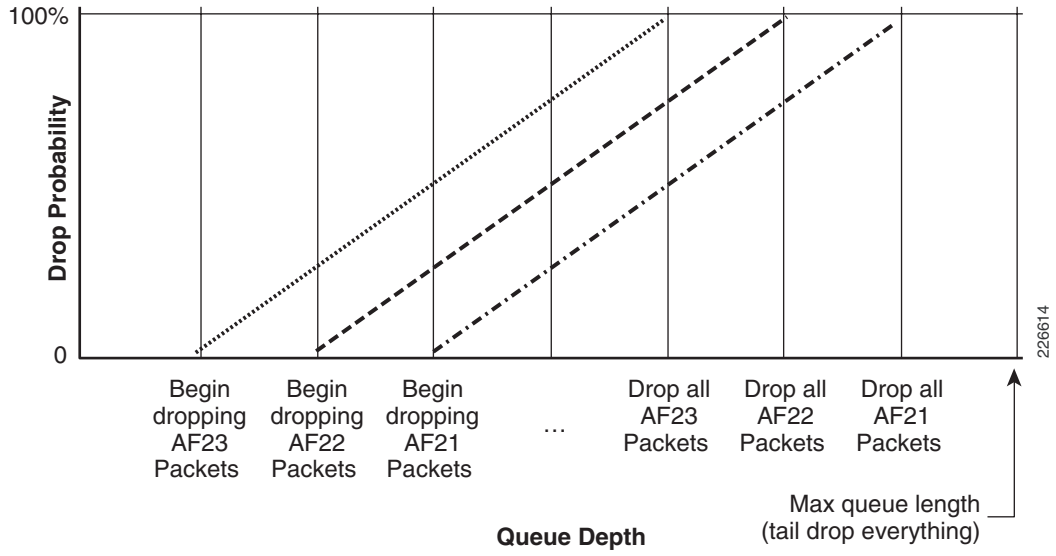
Selective dropping of packets when the queues are filling is referred to as congestion avoidance. Congestion avoidance mechanisms work best with TCP-based applications because selective dropping of packets causes the TCP windowing mechanisms to "throttle-back" and adjust the rate of flows to manageable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

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

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

The operation of DSCP-based WRED is illustrated in Figure 4-20.

Figure 4-20 DSCP-Based WRED Example Operation



Link Efficiency Tools

Link Efficiency Tools are typically relevant only on link speeds ≤768 kbps, and come in two main types:

- Link Fragmentation and Interleaving (LFI) tools—With slow-speed WAN circuits, large data packets take an excessively long time to be placed onto the wire. This delay, called serialization delay, can easily cause a VoIP packet to exceed its delay and/or jitter threshold. There are two LFI tools to mitigate serialization delay on slow speed (≤768 kbps) links, Multilink PPP Link Fragmentation and Interleaving (MLP LFI) and Frame Relay Fragmentation (FRF.12).
- Compression tools—Compression techniques, such as compressed Real-Time Protocol (cRTP), minimize bandwidth requirements and are highly useful on slow links. At 40 bytes total, the header portion of a VoIP packet is relatively large and can account for up to two-thirds or the entire VoIP packet (as in the case of G.729 VoIP). To avoid the unnecessary consumption of available bandwidth, cRTP can be used on a link-by-link basis. cRTP compresses IP/UDP/RTP headers from 40 bytes to between two and five bytes (which results in a bandwidth savings of approximately 66% for G.729 VoIP). However, cRTP is computationally intensive, and therefore returns the best bandwidth-savings value vs. CPU-load on slow speed (≤768 kbps) links.

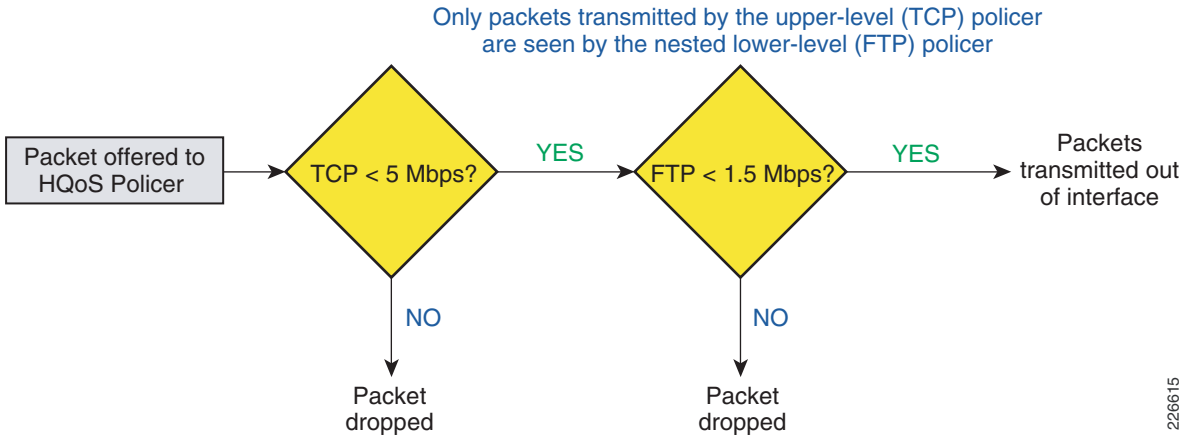
This document is intended to address network designs for today's media networks and, as such, link speeds that are ≤768 kbps are unsuitable in such a context. Therefore, little or no mention is given to link efficiency tools. For networks that still operate at or below 768 kbps, refer to design recommendations within the Enterprise QoS SRND version 3.3 at http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND/QoS-SRND-Book.html

Hierarchical QoS

Cisco IOS MQC-based tools may be combined in a hierarchical fashion, meaning QoS policies may contain other "nested" QoS policies within them. Such policy combinations are commonly referred to as Hierarchal QoS policies or HQoS policies.

Consider a couple of examples where HQoS policies may be useful. In the first case, there may be scenarios where some applications require policing at multiple levels. Specifically, it might be desirable to limit all TCP traffic to 5 Mbps while, at the same time, limiting FTP traffic (which is a subset of TCP traffic) to no more than 1.5 Mbps. To achieve this nested policing requirement, Hierarchical Policing can be used. The policer at the second level in the hierarchy acts on packets transmitted or marked by the policer at the first level, as illustrated in Figure 4-21. Therefore, any packets dropped by the first level are not seen by the second level. Up to three nested levels are supported by the Cisco IOS Hierarchical Policing feature.

Figure 4-21 Hierarchical Policing Policy Example



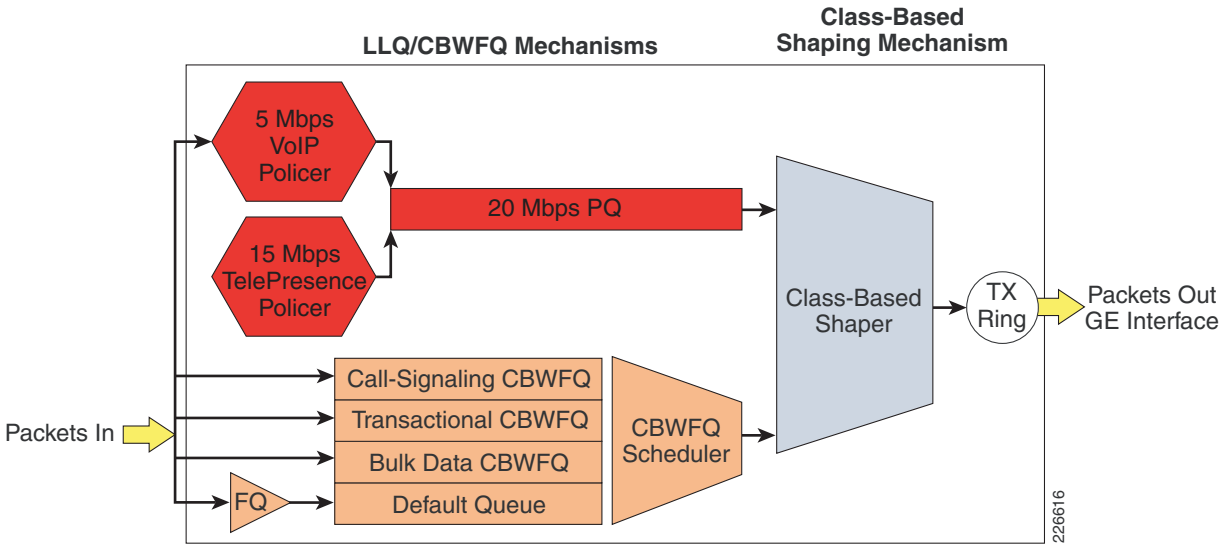
Additionally, it is often useful to combine shaping and queuing policies in a hierarchical manner, particularly over sub-line rate access scenarios. As previously discussed, queuing policies only engage when the physical interface is congested (as is indicated to IOS software by a full Tx-Ring). This means that queuing policies never engage on media that has a contracted sub-line rate of access, whether this media is Frame Relay, ATM, or Ethernet. In such a scenario, queuing can only be achieved at a sub-line rate by introducing a two-part HQoS policy wherein:

- Traffic is shaped to the sub-line rate.
- Traffic is queued according to the LLQ/CBWFQ policies within the sub-line rate.

With such an HQoS policy, it is not the Tx-Ring that signals IOS software to engage LLQ/CBWFQ policies, but rather it is the Class-Based Shaper that triggers software queuing when the shaped rate has been reached.

Consider a practical example in which a service provider offers an enterprise subscriber a GigabitEthernet handoff, but with a (sub-line rate) contract for only 60 Mbps, over which he wants to deploy IP Telephony and TelePresence, as well as applications. Normally, queuing policies only engage on this GE interface when the offered traffic rate exceeds 1000 Mbps. However, the enterprise administrator wants to ensure that traffic within the 60 Mbps contracted rate is properly prioritized prior to the handoff so that both VoIP and TelePresence are given the highest levels of service. Therefore, the administrator configures an HQoS policy, such that the software shapes all traffic to the contracted 60 Mbps rate and attaches a nested LLQ/CBWFQ queuing policy within the shaping policy, such that traffic is properly prioritized within this 60 Mbps sub-line rate. Figure 4-22 illustrates the underlying mechanisms for this HQoS policy.

Figure 4-22 Hierarchical Shaping and Queuing Policy Example



AutoQoS

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

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

In the second release, AutoQoS Enterprise, this feature consists of two configuration phases, completed in the following order:

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

Some may naturally then ask, Why should I read this lengthy and complex QoS design document when I have AutoQoS? It is true that AutoQoS-VoIP is an excellent tool for customers with the objective of enabling QoS for VoIP (only) on their campus and WAN infrastructures. It is also true that AutoQoS-Enterprise is a fine tool for enabling basic branch-router WAN-Edge QoS for voice, video, and multiple classes of data. And as such, customers that have such basic QoS needs and/or do not have the time or desire to do more with QoS, AutoQoS is definitely the way to go.

However, it is important to remember where AutoQoS came from. AutoQoS tools are the result of Cisco QoS feature development coupled with Cisco QoS design guides based on large-scale lab-testing. AutoQoS VoIP is the product of the first QoS design guide (published in 1999). AutoQoS Enterprise is based on the second QoS design guide (published in 2002) and the AutoQoS feature has not been updated since. Therefore, if the business requirements for QoS are quite basic, then—as mentioned—AutoQoS would be an excellent tool to expedite the QoS deployment. If, on the other hand, there are more advanced requirements of QoS—such as those presented in this document—then the configurations presented herein would be recommended over AutoQoS.

QoS Management

Cisco offers a variety of applications to manage quality of service, including

- Cisco QoS Policy Manager (QPM)—QPM supports centralized management of network QoS by providing comprehensive QoS provisioning and monitoring capabilities to deploy, tune, monitor, and optimize the performance characteristics of the network. QPM leverages intelligent network services such as NBAR and other QoS features to identify and monitor networked applications and control their behavior throughout the network.
- Cisco Bandwidth Quality Manager (BQM)—BQM provides end-to-end network service quality monitoring with unique visibility and analysis of traffic, bandwidth, and service quality on IP access networks. BQM can be used to monitor, troubleshoot, and assure end-to-end network performance objectives for converged application traffic. BQM provides micro-level visibility into the network and the network service quality events compromising user experience.
- Cisco Network Analysis Modules (NAM)—Available as Cisco router network modules or as Cisco Catalyst 6500 linecard modules, NAMs can perform extensive voice quality monitoring, intelligent application performance analytics, QoS analysis, and advanced troubleshooting.

Such tools can enable administrators to more efficiently baseline, deploy, monitor, and manage QoS policies over their network infrastructure.

Admission Control Tools

Interactive applications—particularly voice and video applications—often require realtime services from the network. As these resources are finite, they must be managed efficiently and effectively. If the number of flows contending for such priority resources were not limited, then as these resources become oversubscribed, the quality of all realtime flows would degrade—eventually to the point of unusability.

Note Admission Control (AC) is sometimes also referred to as Call Admission Control (CAC); however, as applications evolve, not all applications requiring priority services are call-oriented, and as such AC is a more encompassing designation.

Admission control functionality is most effectively controlled an application-level, such as is the case with Cisco Unified CallManager, which controls VoIP and IP video and/or TelePresence flows. As such, admission control design is not discussed in detail in this document, but will be deferred to application-specific design guides, such as the Cisco Unified Communications design guides and/or the Cisco TelePresence design guides at www.cisco.com/go/designzone.

As discussed previously, media applications are taxing networks as never before. To that end, current admission control tools are not sufficient to make the complex decisions that many collaborative media applications require. Thus, admission control continues to be an field for extended research and development in the coming years, with the goal of developing multi-level admission control solutions, as described below:

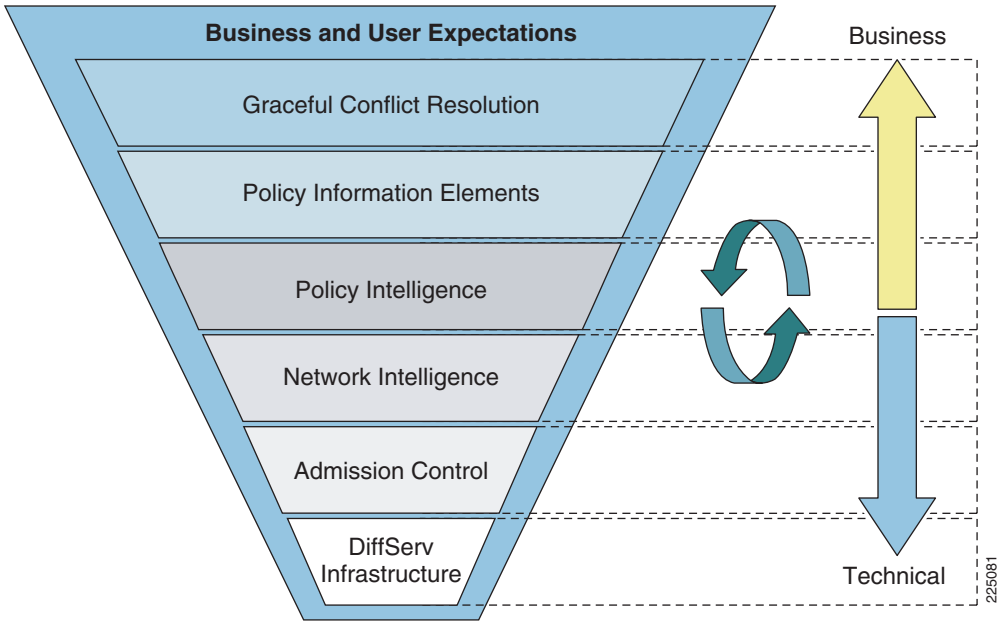
- The first level of admission control is simply to enable mechanisms to protect voice-from-voice and/or video-from-video on a first-come, first-serve basis. This functionality provides a foundation on which higher-level policy-based decisions can be built.
- The second level of admission control factors in dynamic network topology and bandwidth information into a real-time decision of whether or not a media stream should be admitted. These decisions could be made by leveraging intelligent network protocols, such as Resource Reservation Protocol (RSVP).
- The third level of admission control introduces the ability to preempt existing flows in favor of “higher-priority” flows.
- The fourth level of admission control contains policy elements and weights to determine what exactly constitutes a “higher-priority” flow, as defined by the administrative preferences of an organization. Such policy information elements may include—but are not limited to—the following:
  - Scheduled versus ad hoc—Media flows that have been scheduled in advance would likely be granted priority over flows that have been attempted ad hoc.
  - Users and groups—Certain users or user groups may be granted priority for media flows.
  - Number of participants—Multipoint media calls with larger number of participants may be granted priority over calls with fewer participants.
  - External versus internal participants—Media sessions involving external participants, such as customers, may be granted priority over sessions comprised solely of internal participants.
  - Business critical factor—Additional subjective elements may be associated with media streams, such as a business critical factor. For instance, a live company meeting would likely be given a higher business critical factor than a live training session. Similarly, a media call to close a sale or to retain a customer may be granted priority over regular, ongoing calls.

Note It should be emphasized this is not an exhaustive list of policy information elements that could be used for admission control, but rather is merely a sample list of possible policy information elements. Additionally, each of these policy information elements could be assigned administratively-defined weights to yield an overall composite metric to calculate and represent the final admit/deny admission control decision for the stream.

- The fifth level of admission control provides graceful conflict resolution, such that—should preemption of a media flow be required—existing flow users are given a brief message indicating that their flow is about to be preempted (preferably including a brief reason as to why) and a few seconds to make alternate arrangements (as necessary).

A five-level admission control model, deployed over a DiffServ-enabled infrastructure is illustrated in Figure 4-23.

Figure 4-23 Five-Level Admission Control Model Deployed Over a DiffServ Infrastructure



Thus, having laid a foundational context by reviewing QoS technologies, let us turn our attention to Cisco's strategic QoS recommendations for enterprise medianets.

Enterprise Medianet Strategic QoS Recommendations

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

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

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

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

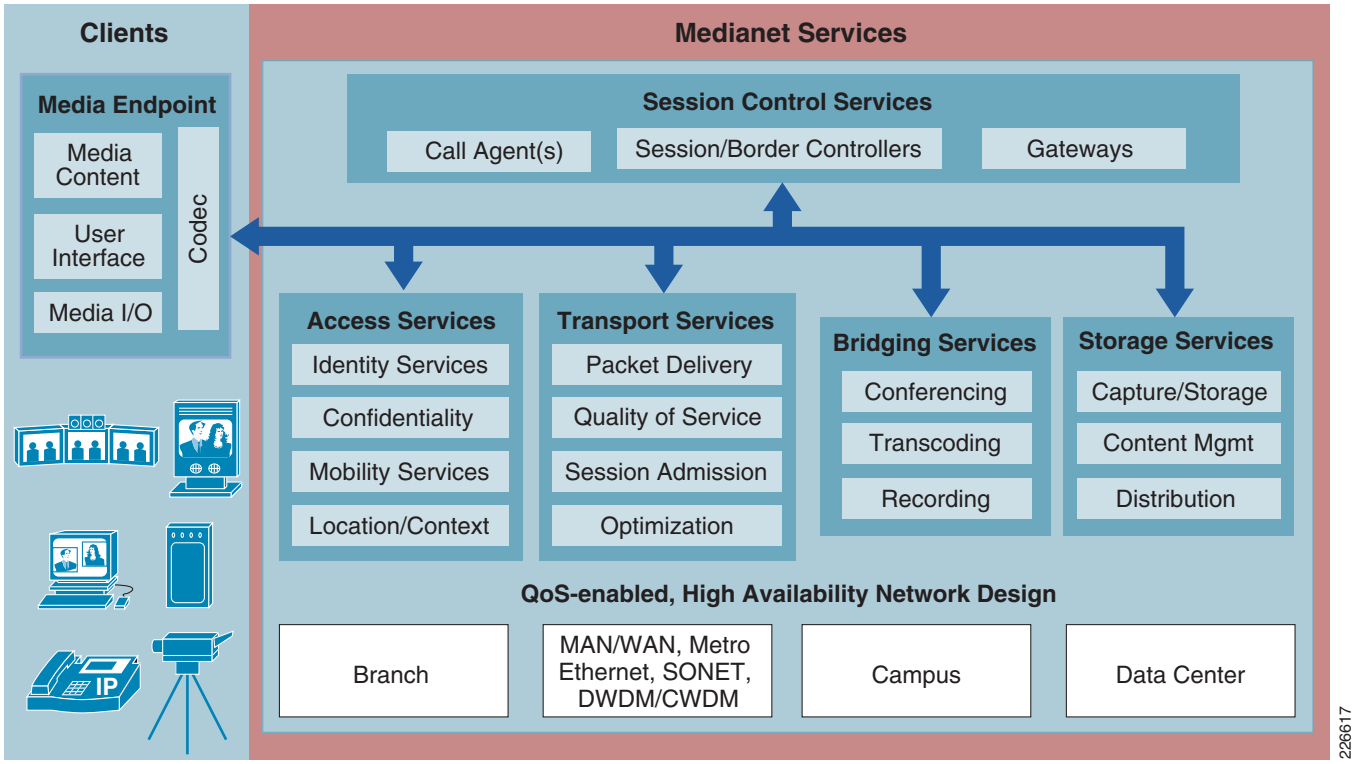
Finally, as the level of corporate intellectual property migrates into stored and interactive media, it is critical to have a strategy to manage the media content, setting and enforcing clear policies, and having the ability to protect intellectual property in secure and managed systems. Just as companies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

Therefore, to meet all these media application requirements, Cisco recommends—not to reengineer networks to support each wave of applications—but rather to utilize an architectural approach, namely a medianet architecture.

Enterprise Medianet Architecture

A medianet is built upon an architecture that supports the different models of media applications and optimizes their delivery, such as those shown in the architectural framework in Figure 4-24.

Figure 4-24 Enterprise Medianet Architectural Framework



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

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

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

With this architectural framework in mind, let us take a closer look at the strategic QoS recommendations for a medianet.

Enterprise Medianet QoS Application Class Recommendations

As mentioned previously, Cisco has slightly modified its implementation of (informational) RFC 4594 (as shown in Figure 4-9). With Admission Control recommendations added to this model, these combined recommendations are summarized in Figure 4-25.

Figure 4-25 Enterprise Medianet QoS Recommendations

Application Class	Per-Hop Behavior	Admission Control	Queuing and Dropping	Media Application Examples
VoIP Telephony	EF	Required	Priority Queue (PQ)	Cisco IP Phones (G.711, G.729)
Broadcast Video	CS5	Required	(Optional) PQ	Cisco IP Video Surveillance/Cisco Enterprise TV
Real-Time Interactive	CS4	Required	(Optional) PQ	Cisco TelePresence
Multimedia Conferencing	AF4	Required	BW Queue + DSCP WRED	Cisco Unified Personal Communicator
Multimedia Streaming	AF3	Recommended	BW Queue + DSCP WRED	Cisco Digital Media System (VoDs)
Network Control	CS6		BW Queue	EIGRP, OSPF, BGP, HSRP, IKE
Signaling	CS3		BW Queue	SCCP, SIP, H.323
Ops/Admin/Mgmt (OAM)	CS2		BW Queue	SNMP, SSH, Syslog
Transactional Data	AF2		BW Queue + DSCP WRED	Cisco WebEx/MeetingPlace/ERP Apps
Bulk Data	AF1		BW Queue + DSCP WRED	E-mail, FTP, Backup Apps, Content Distribution
Best Effort	DF		Default Queue + RED	Default Class
Scavenger	CS1		Min BW Queue	YouTube, iTunes, BitTorrent, Xbox Live

224550

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

- VoIP Telephony
- Broadcast Video
- Realtime Interactive
- Multimedia Conferencing
- Network Control
- Signaling
- Operations, Administration, and Management (OAM)
- Transactional Data and Low-Latency Data
- Bulk Data and High-Throughput Data
- Best Effort
- Scavenger and Low-Priority Data

VoIP Telephony

This service class is intended for VoIP telephony (bearer-only) traffic (VoIP signaling traffic is assigned to the Call Signaling class). Traffic assigned to this class should be marked EF (DSCP 46) and should be admission controlled. This class is provisioned with an Expedited Forwarding Per-Hop Behavior. The EF PHB-defined in RFC 3246-is a strict-priority queuing service and as such, admission to this class should be controlled. Example traffic includes G.711 and G.729a.

Broadcast Video

This service class is intended for broadcast TV, live events, video surveillance flows, and similar "inelastic" streaming media flows ("inelastic" flows refer to flows that are highly drop sensitive and have no retransmission and/or flow-control capabilities). Traffic in this class should be marked Class Selector 5 (CS5/DSCP 40) and may be provisioned with an EF PHB; as such, admission to this class should be controlled (either by an explicit admission control mechanisms or by explicit bandwidth provisioning). Examples traffic includes live Cisco Digital Media System (DMS) streams to desktops or to Cisco Digital Media Players (DMPs), live Cisco Enterprise TV (ETV) streams, and Cisco IP Video Surveillance (IPVS).

Realtime Interactive

This service class is intended for inelastic high-definition interactive video applications and is intended primarily for audio and video components of these applications. Whenever technically possible and administratively feasible, data sub-components of this class can be separated out and assigned to the Transactional Data traffic class. Traffic in this class should be marked CS4 (DSCP 32) and may be provisioned with an EF PHB; as such, admission to this class should be controlled. An example application is Cisco TelePresence.

Multimedia Conferencing

This service class is intended for desktop software multimedia collaboration applications and is intended primarily for audio and video components of these applications. Whenever technically possible and administratively feasible, data sub-components of this class can be separated out and assigned to the Transactional Data traffic class. Traffic in this class should be marked Assured Forwarding Class 4 (AF41/DSCP 34) and should be provisioned with a guaranteed bandwidth queue with DSCP-based Weighted-Random Early Detect (DSCP-WRED) enabled. Admission to this class should be controlled; additionally, traffic in this class may be subject to policing and re-marking. Example applications include Cisco Unified Personal Communicator, Cisco Unified Video Advantage, and the Cisco Unified IP Phone 7985G.

Network Control

This service class is intended for network control plane traffic, which is required for reliable operation of the enterprise network. Traffic in this class should be marked CS6 (DSCP 48) and provisioned with a (moderate, but dedicated) guaranteed bandwidth queue. WRED should not be enabled on this class, as network control traffic should not be dropped (if this class is experiencing drops, then the bandwidth allocated to it should be re-provisioned). Example traffic includes EIGRP, OSPF, BGP, HSRP, IKE, etc.

Signaling

This service class is intended for signaling traffic that supports IP voice and video telephony; essentially, this traffic is control plane traffic for the voice and video telephony infrastructure. Traffic in this class should be marked CS3 (DSCP 24) and provisioned with a (moderate, but dedicated) guaranteed bandwidth queue. WRED should not be enabled on this class, as signaling traffic should not be dropped (if this class is experiencing drops, then the bandwidth allocated to it should be re-provisioned). Example traffic includes SCCP, SIP, H.323, etc.

Operations, Administration, and Management (OAM)

This service class is intended for—as the name implies—network operations, administration, and management traffic. This class is important to the ongoing maintenance and support of the network. Traffic in this class should be marked CS2 (DSCP 16) and provisioned with a (moderate, but dedicated) guaranteed bandwidth queue. WRED should not be enabled on this class, as OAM traffic should not be dropped (if this class is experiencing drops, then the bandwidth allocated to it should be re-provisioned). Example traffic includes SSH, SNMP, Syslog, etc.

Transactional Data and Low-Latency Data

This service class is intended for interactive, "foreground" data applications ("foreground" applications refer to applications from which users are expecting a response—via the network—in order to continue with their tasks. Excessive latency in response times of foreground applications directly impacts user productivity). Traffic in this class should be marked Assured Forwarding Class 2 (AF21 / DSCP 18) and should be provisioned with a dedicated bandwidth queue with DSCP-WRED enabled. This traffic class may be subject to policing and re-marking. Example applications include data components of multimedia collaboration applications, Enterprise Resource Planning (ERP) applications, Customer Relationship Management (CRM) applications, database applications, etc.

Bulk Data and High-Throughput Data

This service class is intended for non-interactive "background" data applications ("background" applications refer to applications from which users are not awaiting a response—via the network—in order to continue with their tasks. Excessive latency in response times of background applications does not directly impact user productivity. Furthermore, as most background applications are TCP-based file-transfers, these applications—if left unchecked—could consume excessive network resources away from more interactive, foreground applications). Traffic in this class should be marked Assured Forwarding Class 1 (AF11/DSCP 10) and should be provisioned with a moderate, but dedicated bandwidth queue with DSCP-WRED enabled. This traffic class may be subject to policing and re-marking. Example applications include E-mail, backup operations, FTP/SFTP transfers, video and content distribution, etc.

Best Effort

This service class is the default class. As only a relative minority of applications are assigned to priority, guaranteed-bandwidth, or even to deferential service classes, the vast majority of applications continue to default to this best effort service class; as such, this default class should be adequately provisioned (a minimum bandwidth recommendation, for this class is 25%). Traffic in this class is marked Default Forwarding (DF or DSCP 0) and should be provisioned with a dedicated queue. WRED is recommended to be enabled on this class. Although, since all the traffic in this class is marked to the same "weight" (of DSCP 0), the congestion avoidance mechanism is essentially Random Early Detect (RED).

Scavenger and Low-Priority Data

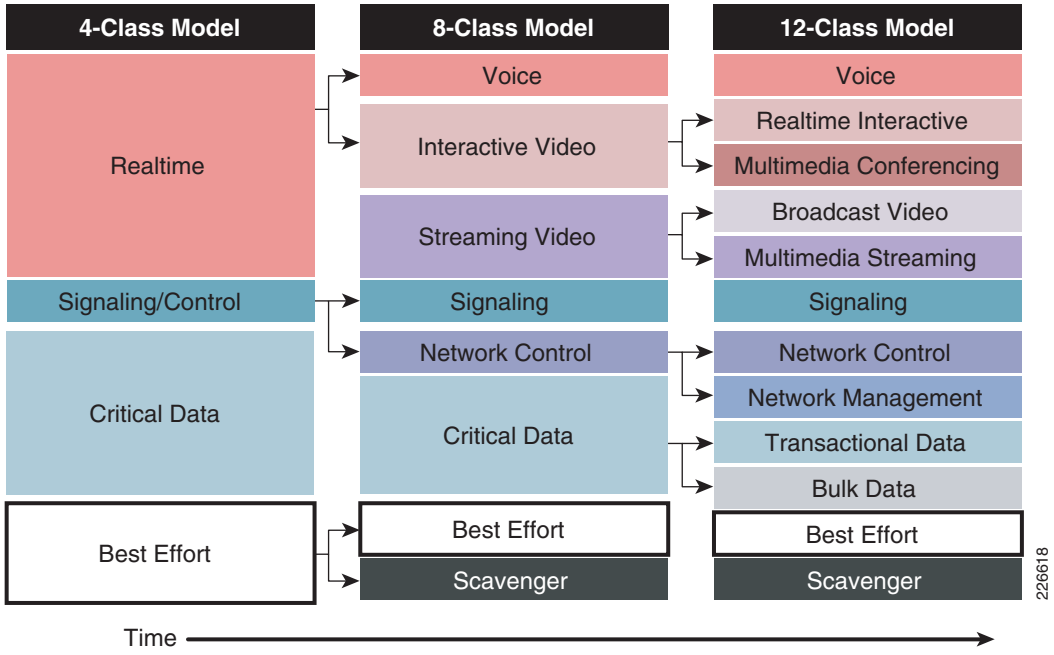
This service class is intended for non-business related traffic flows, such as data or media applications that are entertainment-oriented. The approach of a less-than best effort service class for non-business applications (as opposed to shutting these down entirely) has proven to be a popular, political compromise. These applications are permitted on enterprise networks, as long as resources are always available for business-critical media applications. However, as soon the network experiences congestion, this class is the first to be penalized and aggressively dropped. Furthermore, the scavenger class can be

utilized as part of an effective strategy for DoS and worm attack mitigation (discussed later in this chapter). Traffic in this class should be marked CS1 (DSCP 8) and should be provisioned with a minimal bandwidth queue that is the first to starve should network congestion occur. Example traffic includes YouTube, Xbox Live/360 Movies, iTunes, BitTorrent, etc.

Media Application Class Expansion

While there are merits to adopting a 12-class model, as outlined in the previous section, Cisco recognizes that not all enterprises are ready to do so, whether this be due to business reasons, technical constraints, or other reasons. Therefore, rather than considering these medianet QoS recommendations as an all-or-nothing approach, Cisco recommends considering a phased approach to media application class expansion, as illustrated in Figure 4-26.

Figure 4-26 Media Application Class Expansion



Utilizing such a phased approach to application class expansion, enterprise administrators can incrementally implement QoS policies across their infrastructures in a progressive manner, inline with their business needs and technical constraints. Familiarity with this enterprise medianet QoS model can assist in the smooth expansion of QoS policies to support additional media applications as future requirements arise. Nonetheless, at the time of QoS deployment, the enterprise needs to clearly define their business objectives with QoS, which correspondingly determines how many traffic classes will be required at each phase of deployment.

Cisco QoS Best Practices

With an overall application PHB strategy in place, end-to-end QoS policies can be designed for each device and interface, as determined by their roles in the network infrastructure. These are detailed in the various PIN-specific QoS design chapters that follow. However, because the Cisco QoS toolset provides many QoS design and deployment options, a few succinct design principles can help simplify strategic QoS deployments.

Hardware versus Software QoS

A fundamental QoS design principle is to **always enable QoS policies in hardware—rather than software—whenever a choice exists**. Cisco IOS routers perform QoS in software, which places incremental loads on the CPU, depending on the complexity and functionality of the policy. Cisco Catalyst switches, on the other hand, perform QoS in dedicated hardware ASICS on Ethernet-based ports and as such do not tax their main CPUs to administer QoS policies. This allows complex policies to be applied at line rates at even Gigabit or Ten-Gigabit speeds.

Classification and Marking Best Practices

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

In general, it is not recommended to trust markings that can be set by users on their PCs or other similar devices, because users can easily abuse provisioned QoS policies if permitted to mark their own traffic. For example, if an EF PHB has been provisioned over the network, a PC user can easily configure all their traffic to be marked to EF, thus hijacking network priority queues to service non-realtime traffic. Such abuse could easily ruin the service quality of realtime applications throughout the enterprise. On the other hand, if enterprise controls are in place that centrally administer PC QoS markings, then it may be possible and advantageous to trust these.

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

As the line between enterprises and service providers continues to blur and the need for interoperability and complementary QoS markings is critical, you should **follow standards-based DSCP PHB markings to ensure interoperability and future expansion**. Because the enterprise medianet marking recommendations are standards-based—as has been previously discussed—enterprises can easily adopt these markings to interface with service provider classes of service. Network mergers—whether the result of acquisitions, mergers, or strategic alliances—are also easier to manage when using standards-based DSCP markings.

Policing and Markdown Best Practices

There is little reason to forward unwanted traffic only to police and drop it at a subsequent node, especially when the unwanted traffic is the result of DoS or worm attacks. Furthermore, the overwhelming volume of traffic that such attacks can create can cause network outages by driving

network device processors to their maximum levels. Therefore, it is recommended to **police traffic flows as close to their sources as possible**. This principle applies also to legitimate flows, as worm-generated traffic can masquerade under legitimate, well-known TCP/UDP ports and cause extreme amounts of traffic to be poured onto the network infrastructure. Such excesses should be monitored at the source and marked down appropriately.

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

Queuing and Dropping Best Practices

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

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

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

Additional queuing recommendations for these classes are discussed next.

Strict-Priority Queuing Recommendations—The 33 Percent LLQ Rule

The Realtime or Strict Priority class corresponds to the RFC 3246 EF PHB. The amount of bandwidth assigned to the realtime queuing class is variable. However, if the majority of bandwidth is provisioned with strict priority queuing (which is effectively a FIFO queue), then the overall effect is a dampening of QoS functionality, both for latency and jitter sensitive realtime applications (contending with each other within the FIFO priority queue) and also for non-realtime applications (as these may periodically receive wild bandwidth allocation fluctuations, depending on the instantaneous amount of traffic being serviced by the priority queue). Remember the goal of convergence is to enable voice, video, and data applications to transparently co-exist on a single IP network. When realtime applications dominate a link, then non-realtime applications fluctuate significantly in their response times, destroying the transparency of the converged network.

For example, consider a (45 Mbps) DS3 link configured to support two TelePresence (CTS-3000) calls with an EF PHB service. Assuming that both systems are configured to support full high definition, each such call requires 15 Mbps of strict-priority queuing. Prior to TelePresence calls being placed, non-realtime applications have access to 100% of the bandwidth on the link (to simplify the example, assume there are no other realtime applications on this link). However, once these TelePresence calls are established, all non-realtime applications would suddenly be contending for less than 33% of the link.

TCP windowing would take effect and many applications hang, time-out, or become stuck in a non-responsive state, which usually translates into users calling the IT help desk complaining about the network (which happens to be functioning properly, albeit in a poorly-configured manner).

To obviate such scenarios, Cisco Technical Marketing has done extensive testing and has found that a significant decrease in non-realtime application response times occurs when realtime traffic exceeds one-third of link bandwidth capacity. Extensive testing and customer deployments have shown that a general best queuing practice is to **limit the amount of strict priority queuing to 33% of link bandwidth capacity**. This strict priority queuing rule is a conservative and safe design ratio for merging realtime applications with data applications.

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

It is vitally important to understand that **this strict priority queuing rule is simply a best practice design recommendation and is not a mandate**. There may be cases where specific business objectives cannot be met while holding to this recommendation. In such cases, enterprises must provision according to their detailed requirements and constraints. However, it is important to recognize the tradeoffs involved with over-provisioning strict priority traffic and its negative performance impact both on other realtime flows and also on non-realtime-application response times.

And finally, **any traffic assigned to a strict-priority queue should be governed by an admission control mechanism**.

Best Effort Queuing Recommendation

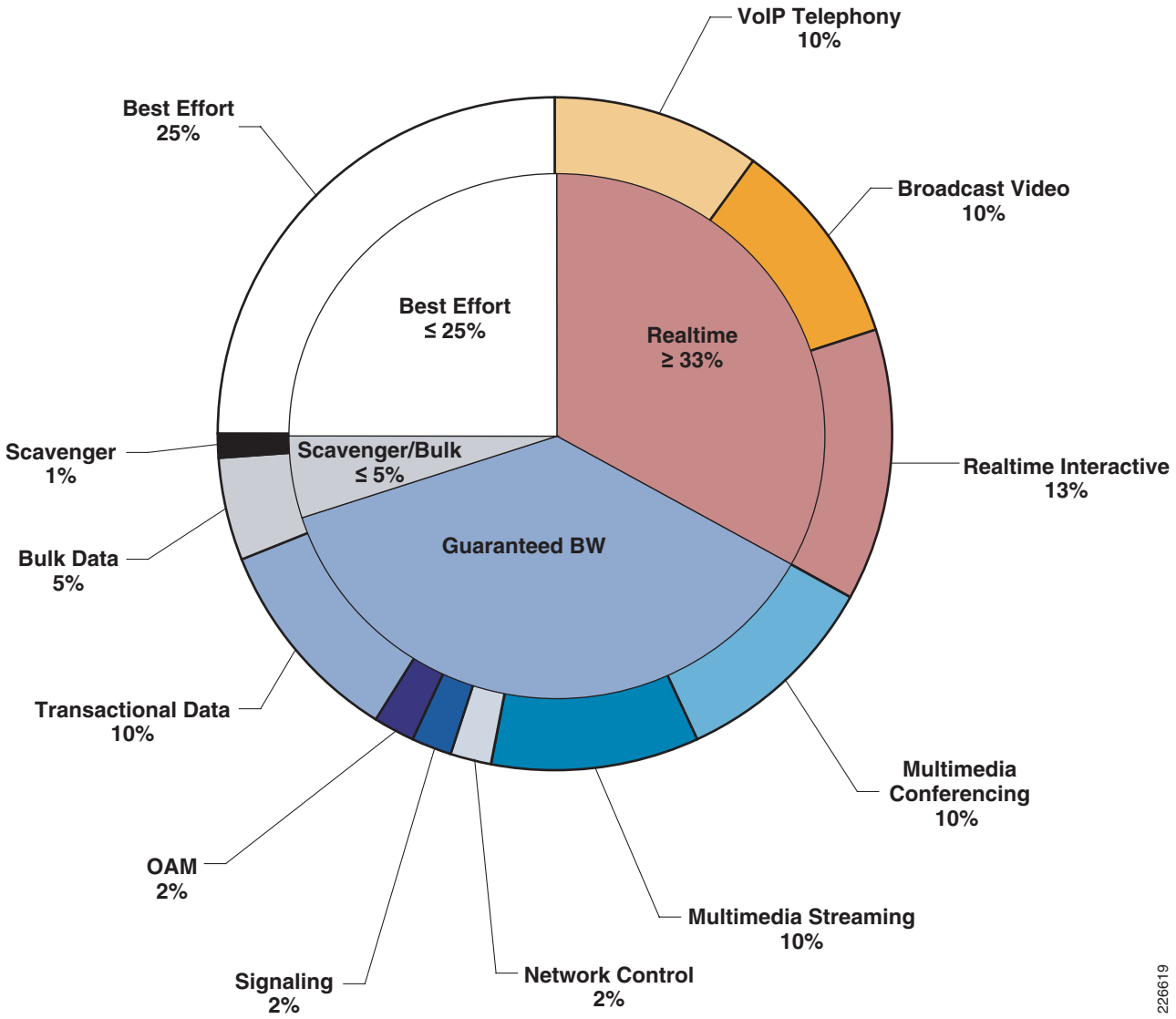
The Best Effort class is the default class for all traffic that has not been explicitly assigned to another application-class queue. Only if an application has been selected for preferential/deferential treatment is it removed from the default class. Because most enterprises have several thousand applications running over their networks, adequate bandwidth must be provisioned for this class as a whole in order to handle the sheer number and volume of applications that default to it. Therefore, it is recommended to **reserve at least 25 percent of link bandwidth for the default Best Effort class**.

Scavenger Class Queuing Recommendations

Whenever Scavenger queuing class is enabled, it should be assigned a minimal amount of bandwidth, such as 1% (or whatever the minimal bandwidth allocation that the platform supports). On some platforms, queuing distinctions between Bulk Data and Scavenger traffic flows cannot be made, either because queuing assignments are determined by CoS values (and both of these application classes share the same CoS value of 1) or because only a limited amount of hardware queues exist, precluding the use of separate dedicated queues for each of these two classes. In such cases, the Scavenger/Bulk queue can be assigned moderate amount of bandwidth, such as 5%.

These queuing rules are summarized in Figure 4-27, where the inner pie chart represents a hardware or service provider queuing model that is limited to four queues and the outer pie chart represents a corresponding, more granular queuing model that is not bound by such constraints.

Figure 4-27 Compatible 4-Class and 12-Class Medianet Queuing Models



QoS for Security Best Practices

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

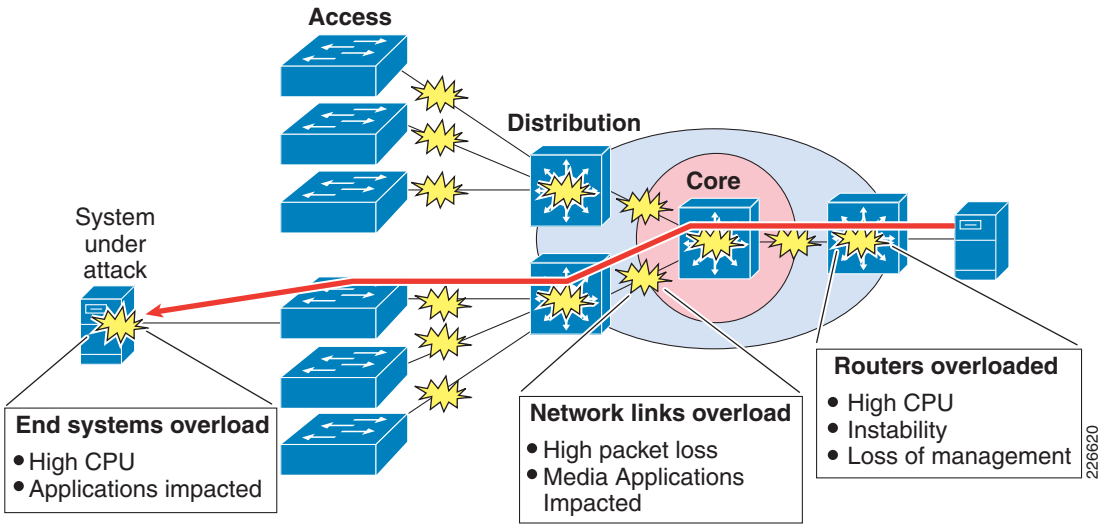
There are two main classes of DoS attacks:

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

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

In contrast, worms exploit security vulnerabilities in their targets and disguisedly carry harmful payloads that usually include a self-propagating mechanism. Network infrastructure usually is not the direct target of a worm attack, but can become collateral damage as worms exponentially self-propagate. The rapidly multiplying volume of traffic flows eventually drowns the CPU/hardware resources of routers and switches in their paths, indirectly causing Denial of Service to legitimate traffic flows, as shown in Figure 4-28.

Figure 4-28 Direct and Indirect Collateral Damage from DoS/Worm Attacks



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

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

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

A *proactive* approach to mitigating DoS/worm attacks within enterprise networks is to have control plane policing and data plane policing policies in place within the infrastructure which immediately respond to out-of-profile network behavior indicative of DoS or worm attacks. Control plane policing serves to protect the CPU of network devices—such as switches and routers—from becoming bogged down with interruption-handling and thus not having enough cycles to forward traffic. Data plane policing—also referred to as Scavenger-class QoS—serves to protect link bandwidth from being consumed by forwarding DoS/worm traffic to the point of having no room to service legitimate, in-profile flows.

Control Plane Policing

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

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

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

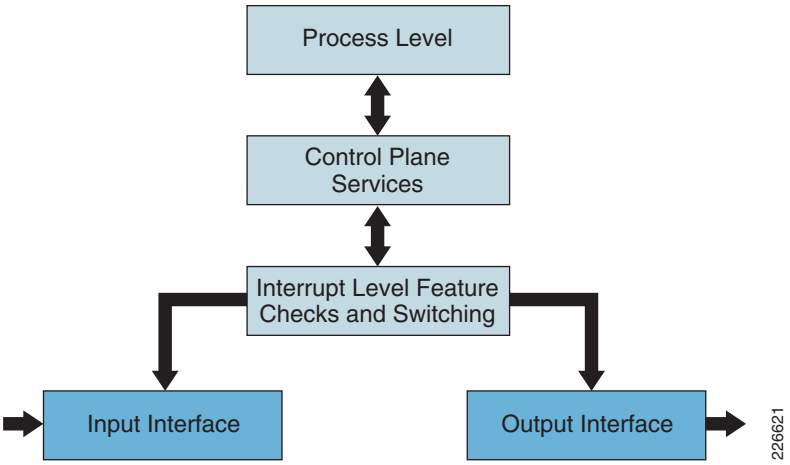
Because the route processor is critical to network operations, any service disruption to the route processor or the control and management planes can result in business-impacting network outages. A DoS attack targeting the route processor, which can be perpetrated either inadvertently or maliciously,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

- High route processor CPU utilization (near 100%)
- Loss of line protocol keepalives and routing protocol updates, leading to route flaps and major network transitions
- Interactive sessions via the Command Line Interface (CLI) are slow or completely unresponsive due to high CPU utilization
- Route processor resource exhaustion—resources such as memory and buffers are unavailable for legitimate IP data packets
- Packet queue backup, which leads to indiscriminate drops (or drops due to lack of buffer resources) of other incoming packets

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

Figure 4-29 illustrates the flow of packets from various interfaces. Packets destined to the control plane are subject to control plane policy checking, as depicted by the control plane services block.

Figure 4-29 Packet Flow Within a Switch/Router



By protecting the route processor, CPP/CoPP helps ensure router and network stability during an attack. For this reason, **a best practice recommendation is to deploy CPP/CoPP as a key protection mechanism on all routers and switches that support this feature**.

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

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

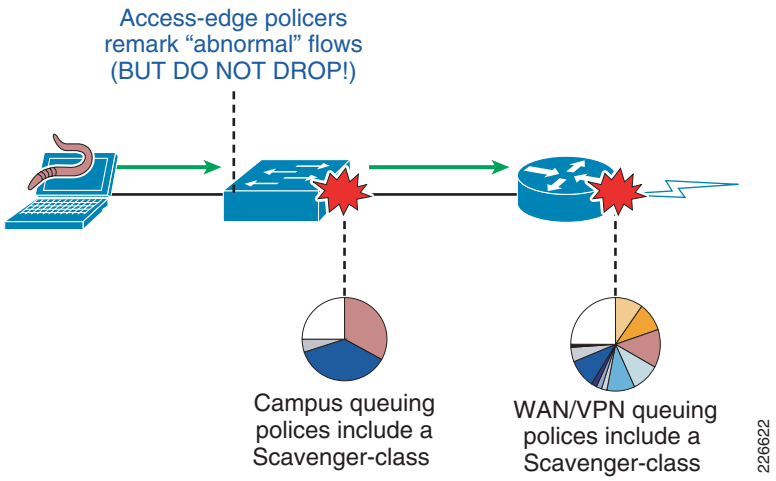
Data Plane Policing/Scavenger-Class QoS

The logic applied to protecting the control plane can also be applied to the data plane. Data plane policing has two components:

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

These two components of data plane policing/Scavenger-class QoS are illustrated in Figure 4-30.

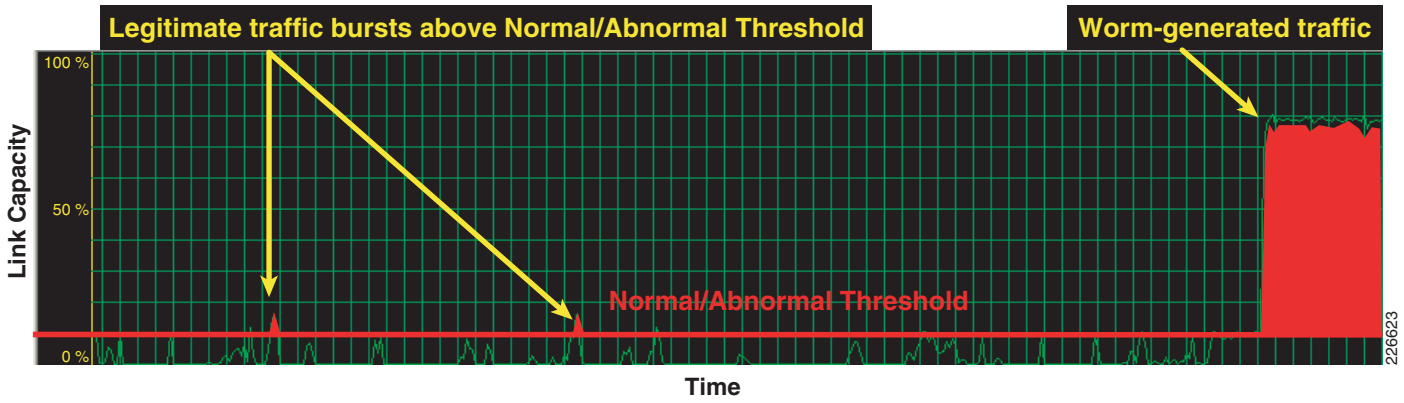
Figure 4-30 Data Plane Policing/Scavenger-class QoS Components



Most endpoint devices have fairly predictable traffic patterns and, as such, can have metering policers to identify “normal” flows (the volume of traffic that represents 95% of the typically-generated traffic rates for the endpoint device) vs. “abnormal” flows (the remainder). For instance, it would be “abnormal” for a port that supposedly connects to an IP phone to receive traffic in excess of 128 kbps. Similarly, it would be “abnormal” for a port that supposedly connects to a Cisco TelePresence system to receive traffic in excess of 20 Mbps. Both scenarios would be indicative of network abuse—either intentional or inadvertent. Endpoint PCs also have traffic patterns that can be fairly accurately baselined with statistical analysis.

For example, for users of Windows-based systems, the Windows Task Manager (which can be selected by simultaneously pressing CTRL-ALT-DEL) can graphically display networking statistics (available from the networking tab). Most users are generally surprised at how low the average network utilization rates of PCs are during everyday use, as compared to their link speed capacities. Such a graphical display of network utilization is shown in Figure 4-31, where the radical and distinctive difference in network utilization rates after worm-infection is highlighted.

Figure 4-31 Sample PC Network Utilization Rates—Before and After Infection by a Worm



These access edge metering policers are relatively unintelligent. They do not match specific network characteristics of specific types of attacks, but simply meter traffic volumes and respond to abnormally high volumes as close to the source as possible. The simplicity of this approach negates the need for the policers to be programmed with knowledge of the specific details of how the attack is being generated or propagated. It is precisely this unintelligence of such access layer metering policers that allow them

to maintain relevancy as worms mutate and become more complex. The policers do not care how the traffic was generated or what it looks like; they care only how much traffic is being put onto the wire. Therefore, they continue to police even advanced worms that continually change their tactics of traffic-generation.

For example, in most enterprises it is quite abnormal (within a 95% statistical confidence interval) for PCs to generate sustained traffic in excess of 5% of link capacity. In the case of a GigabitEthernet access switch port, this means that it would be unusual in most organizations for an end user PC to generate more than 50 Mbps of uplink traffic on a sustained basis.

Note It is important to recognize that this value (5%) for normal endpoint utilization by PC endpoints is just an example value. This value would likely vary from enterprise to enterprise, as well as within a given enterprise (such as by departmental functions).

It is very important to recognize that what is being recommended by data plane policing/Scavenger class QoS is not to police all traffic to 50 Mbps and automatically drop the excess. Should that be the case, there would not be much reason to deploy GigabitEthernet switch ports to endpoint devices But rather, these campus access-layer policers do not drop traffic at all; they only perform remarking (if traffic rates appear abnormal). These policers are coupled with queuing polices on all network nodes that include a deferential service class for traffic marked as Scavenger (CS1). Queuing policies only engage when links are congested; as such, if links capacity exists, then traffic is never dropped. It is only in scenarios where offered traffic flows exceed link capacity—forcing queuing polices to engage and queuing buffers to fill to capacity—that drops may occur. In such scenarios, dropping can either occur indiscriminately (on a last-come-first-dropped basis) or with a degree of intelligence (as would be the case if abnormal traffic flows were previously identified).

Let's illustrate how this might work for both legitimate excess traffic and also the case of illegitimate excess traffic resulting from a DoS or worm attack.

In the former case, assume that the PC generates over 50 Mbps of traffic, perhaps because of a large file transfer or backup. Congestion (under normal operating conditions) is rarely if ever experienced within the campus because there is generally abundant capacity to carry the traffic. Uplinks to the distribution and core layers of the campus network are typically GigabitEthernet or Ten Gigabit Ethernet, which would require 1,000 or 10,000 Mbps of traffic (respectively) from the access layer switch to congest. If the traffic is destined to the far side of a WAN/VPN link, queuing and dropping typically occurs even without the access layer policer, because of the bottleneck caused by the typical campus-to-WAN/VPN speed mismatch. In such a case, the TCP sliding windows mechanism would eventually find an optimal speed (under 50 Mbps) for the file transfer. Access layer policers that markdown out-of-profile traffic to Scavenger (CS1) would thus not affect legitimate traffic, aside from the obvious remarking. *No reordering or dropping would occur on such flows as a result of these data plane policers that would not have occurred anyway.*

In the latter case, the effect of access layer policers on traffic caused by DoS or worm attacks is quite different. As hosts become infected and traffic volumes multiply, congestion may be experienced even within the campus. For example, if just 11 end user PCs on a single access switch begin spawning worm flows to their maximum GigabitEthernet link capacities, even Ten Gigabit Ethernet uplinks/core links will congest and queuing and dropping policies will engage. At this point, VoIP and media applications, and even best effort applications, would gain priority over worm-generated traffic (as Scavenger traffic would be dropped the most aggressively). Furthermore, network devices would remain accessible for administration of the patches/plugs/ACLs/NBAR policies required to fully neutralize the specific attack. WAN/VPN links would also be similarly protected, which is a huge advantage, as generally WAN/VPN links are the first to be overwhelmed by DoS/worm attacks. Scavenger class policies thus *significantly mitigate network traffic generated by DoS or worm attacks.*

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

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

A final word on this subject—it is important to recognize the distinction between mitigating an attack and preventing it entirely. **Control plane policing and data plane policing policies do not guarantee that no DoS or worm attacks will ever happen, but serve only to reduce the risk and impact that such attacks could have on the network infrastructure. Therefore, it is vital to overlay a comprehensive security strategy over the QoS-enabled network infrastructure**.

Summary

This chapter began by discussing the reasons driving the QoS design updates presented in this document by examining three sets of drivers behind QoS design evolution, including:

- New applications and business requirements
- New industry guidance and best practices
- New platforms and technologies

Business drivers—including the evolution of video applications, the phenomena of social networking, the convergence within media applications, the globalization of the workforce, and the pressures to be "green"—were examined to determine how these impact new QoS designs over enterprise media networks. Next, developments in industry standards and best practices—with particular emphasis on RFC 4594 Configuration Guidelines for DiffServ Classes—were discussed, as were developments in QoS technologies and their respective impacts on QoS design.

Cisco's QoS toolset was overviewed to provide a foundational context for the strategic best practices that followed. Classification and marking tools, policing and markdown tools, shaping, queuing, and dropping tools were all reviewed, as were AutoQoS and QoS management tools.

An enterprise medianet architecture was then presented, along with strategic QoS design recommendations. These recommendations included an RFC 4594-based application class model and an application expansion class model (for enterprises not yet ready to deploy a 12-class QoS model). Additionally, QoS best practices for classification, marking, policing, and queuing were presented, including:

- Always deploy QoS in hardware (over software) whenever possible.
- Mark as close to the source as possible with standards-based DSCP values.
- Police as close to the source as possible.
- Markdown according to standards-based rules.
- Deploy queuing policies on all network nodes.

- Optimally assign each medianet class a dedicated queue.
- Limit strict-priority queuing to 33% of link-capacity whenever possible.
- Provision at least 25% of a link's capacity for best effort applications.
- Provision a minimal queue (such as 1%) for the Scavenger applications class.
- Enable control plane policing on platforms that support this feature.
- Deploy data plane policing/Scavenger class QoS polices whenever possible.

These strategic design recommendations will serve to make the PIN-specific designs that follow more cohesive, complementary, and consistent.

References

White Papers

- Cisco Visual Networking Index—Forecast and Methodology, 2007-2012
http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html
- Approaching the Zettabyte Era
http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481374_ns827_Networking_Solutions_White_Paper.html
- Cisco Enterprise QoS Solution Reference Design Guide, version 3.3
http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND/QoS-SRND-Book.html
- Overview of a Medianet Architecture
http://www.cisco.com/en/US/docs/solutions/Enterprise/Video/vrn.html

IETF RFCs

- RFC 791 Internet Protocol
http://www.ietf.org/rfc/rfc791
- RFC 2474 Definition of the Differentiated Services Field
http://www.ietf.org/rfc/rfc2474
- RFC 2597 Assured Forwarding PHB Group
http://www.ietf.org/rfc/rfc2597
- RFC 3246 An Expedited Forwarding PHB
http://www.ietf.org/rfc/rfc3246
- RFC 3662 A Lower Effort Per-Domain Behavior for Differentiated Services
http://www.ietf.org/rfc/rfc3662
- RFC 4594 Configuration Guidelines for DiffServ Service Classes
http://www.ietf.org/rfc/rfc4594
- RFC 5187 [Draft] Aggregation of Diffserv Service Classes
http://tools.ietf.org/html/rfc5127

Cisco Documentation

- Cisco IOS Quality of Service Solutions Configuration Guide, Release 12.4
  http://www.cisco.com/en/US/docs/ios/qos/configuration/guide/12_4/qos_12_4_book.html