This chapter provides an overview of the QoS design and deployment process. This process requires business-level objectives of the QoS implementation to be defined clearly and for the service-level requirements of applications to be assigned preferential or deferential treatment so that they can be analyzed.

These enterprise applications with unique QoS requirements are discussed in this chapter:

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

Additionally, key QoS design and deployment best practices that can simplify and expedite QoS implementations are presented, including these:

- Classification and marking principles
- Policing and markdown principles
- Queuing and dropping principles
- DoS and worm mitigation principles
- Deployment principles

QoS Design Overview

More than just a working knowledge of QoS tools and syntax is needed to deploy end-to-end QoS in a holistic manner. First, it is vital to understand the service-level requirements of the various applications that require preferential (or deferential) treatment within the network. Additionally, a number of QoS design principles that extensive lab testing and customer deployments have helped shape can streamline a QoS deployment and increase the overall cohesiveness of service levels across multiple platforms.

This chapter overviews the QoS requirements of VoIP, Video (both Interactive-Video and Streaming-Video), and multiple classes of data. Within this discussion, the QoS requirements of the control plane (routing and management traffic) are considered. The Scavenger class is examined in more detail, and a strategy for mitigating DoS and worm attacks is presented.

Next, QoS design principles relating to classification, marking, policing, queuing, and deployment are discussed. These serve as guiding best practices in the design chapters to follow.

QoS Requirements of VoIP

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

Voice (Bearer Traffic)

The following list summarizes the key QoS requirements and recommendations for voice (bearer traffic):

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

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

Loss

Loss causes voice clipping and skips. Packet loss concealment (PLC) is a technique used to mask the effects of lost or discarded VoIP packets. The method of PLC used depends upon the type of codec. A simple method used by waveform codecs such as G.711 (PLC for G.711 is defined in G.711 Appendix I) is to replay the last received sample with increasing attenuation at each repeat; the waveform does not change much from one sample to the next. This technique can be effective at concealing the loss of up to 20 ms of samples.

The packetization interval determines the size of samples contained within a single packet. Assuming a 20-ms (default) packetization interval, the loss of two or more consecutive packets results in a noticeable degradation of voice quality. Therefore, assuming a random distribution of drops within a single voice flow, a drop rate of 1 percent in a voice stream would result in a loss that could not be concealed every 3 minutes, on average. A 0.25 percent drop rate would result in a loss that could not be concealed once every 53 minutes, on average.

NOTE A decision to use a 30-ms packetization interval, for a given probability of packet loss, could result in worse perceived call quality than for 20 ms because PLC could not effectively conceal the loss of a single packet.

Low-bit-rate, frame-based codecs, such as G.729 and G.723, use more sophisticated PLC techniques that can conceal up to 30 to 40 ms of loss with "tolerable" quality when the available history used for the interpolation is still relevant.

With frame-based codecs, the packetization interval determines the number of frames carried in a single packet. As with waveform-based codecs, if the packetization interval is greater than the loss that the PLC algorithm can interpolate for, PLC cannot effectively conceal the loss of a single packet.

VoIP networks typically are designed for very close to 0 percent VoIP packet loss, with the only actual packet loss being due to L2 bit errors or network failures.
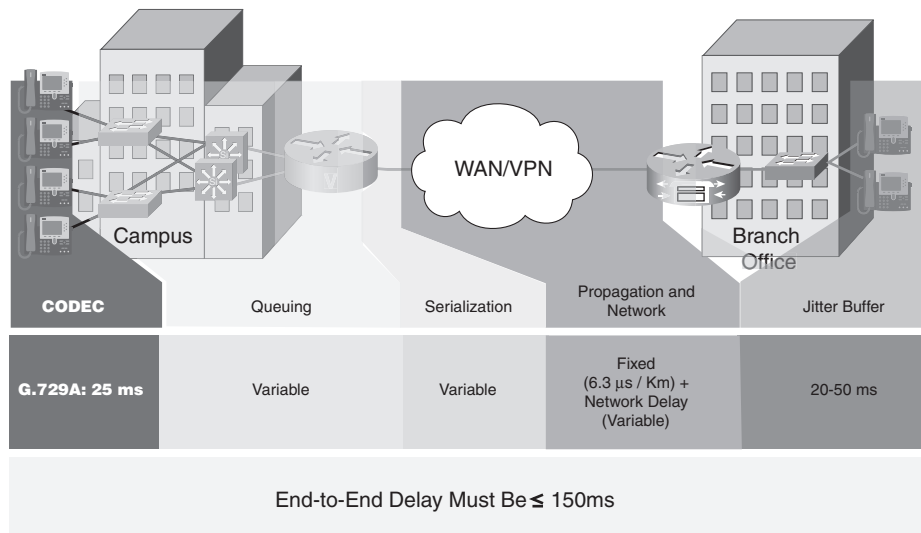
Latency

Latency can cause voice quality degradation if it is excessive. The goal commonly used in designing networks to support VoIP is the target specified by ITU standard G.114 (which, incidentally, is currently under revision): This states that 150 ms of one-way, end-to-end (from mouth to ear) delay ensures user satisfaction for telephony applications. A design

should apportion this budget to the various components of network delay (propagation delay through the backbone, scheduling delay because of congestion, and access link serialization delay) and service delay (because of VoIP gateway codec and dejitter buffer).

Figure 2-1 illustrates these various elements of VoIP latency (and jitter because some delay elements are variable).

Figure 2-1   *Elements Affecting VoIP Latency and Jitter*

CODEC	Queuing	Serialization	Propagation and Network	Jitter Buffer
G.729A: 25 ms	Variable	Variable	Fixed (6.3 µs / Km) + Network Delay (Variable)	20-50 ms

End-to-End Delay Must Be ≤ 150ms

If the end-to-end voice delay becomes too long, the conversation begins to sound like two parties talking over a satellite link or even a CB radio. The ITU G.114 states that a 150-ms one-way (mouth-to-ear) delay budget is acceptable for high voice quality, but lab testing has shown that there is a negligible difference in voice quality mean opinion scores (MOS) using networks built with 200-ms delay budgets. Thus, Cisco recommends designing to the ITU standard of 150 ms. If constraints exist and this delay target cannot be met, the delay boundary can be extended to 200 ms without significant impact on voice quality.

NOTE   Certain organizations might view higher delays as acceptable, but the corresponding reduction in VoIP quality must be taken into account when making such design decisions.

Jitter

Jitter buffers (also known as playout buffers) are used to change asynchronous packet arrivals into a synchronous stream by turning variable network delays into constant delays at the destination end systems. The role of the jitter buffer is to trade off between delay and the probability of interrupted playout because of late packets. Late or out-of-order packets are discarded.

If the jitter buffer is set either arbitrarily large or arbitrarily small, it imposes unnecessary constraints on the characteristics of the network. A jitter buffer set too large adds to the end-to-end delay, meaning that less delay budget is available for the network; hence, the network needs to support a tighter delay target than practically necessary. If a jitter buffer is too small to accommodate the network jitter, buffer underflows or overflows can occur. In an underflow, the buffer is empty when the codec needs to play out a sample. In an overflow, the jitter buffer is already full and another packet arrives; that next packet cannot be enqueued in the jitter buffer. Both jitter buffer underflows and overflows cause voice quality degradation.

Adaptive jitter buffers aim to overcome these issues by dynamically tuning the jitter buffer size to the lowest acceptable value. Well-designed adaptive jitter buffer algorithms should not impose any unnecessary constraints on the network design by doing the following:

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

- Slowly decreasing the jitter buffer size when the measured jitter is less than the current jitter buffer size

- Using PLC to interpolate for the loss of a packet on a jitter buffer underflow

When such adaptive jitter buffers are used—in theory—you can "engineer out" explicit considerations of jitter by accounting for worst-case per-hop delays. Advanced formulas can be used to arrive at network-specific design recommendations for jitter (based on maximum and minimum per-hop delays). Alternatively, because extensive lab testing has shown that voice quality degrades significantly when jitter consistently exceeds 30 ms, this 30 ms value can be used as a jitter target.

Because of its strict service-level requirements, VoIP is well suited to the expedited forwarding per-hop behavior, defined in RFC 3246 (formerly RFC 2598). Therefore, it should be marked to DSCP EF (46) and assigned strict-priority servicing at each node, regardless of whether such servicing is done in hardware (as in Catalyst switches through 1PxQyT queuing, discussed in more detail in Chapter 10, "Catalyst QoS Tools") or in software (as in Cisco IOS routers through LLQ, discussed in more detail in Chapter 5, "Congestion-Management Tools").

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

Table 2-1 details the bandwidth per VoIP flow (both G.711 and G.729) at a default packetization rate of 50 packets per second (pps) and at a custom packetization rate of 33 pps. This does not include Layer 2 overhead and does not take into account any possible compression schemes, such as Compressed Real-Time Transport Protocol (cRTP, discussed in detail in Chapter 7, "Link-Specific Tools").

For example, assume a G.711 VoIP codec at the default packetization rate (50 pps). A new VoIP packet is generated every 20 ms (1 second / 50 pps). The payload of each VoIP packet is 160 bytes; with the IP, UDP, and RTP headers (20 + 8 + 12 bytes, respectively) included, this packet become 200 bytes in length. Converting bits to bytes requires multiplying by 8 and yields 1600 bps per packet. When multiplied by the total number of packets per second (50 pps), this arrives at the Layer 3 bandwidth requirement for uncompressed G.711 VoIP: 80 kbps. This example calculation corresponds to the first row of Table 2-1.

Table 2-1   *Voice Bandwidth (Without Layer 2 Overhead)*

Bandwidth Consumption	Packetization Interval	Voice Payload in Bytes	Packets Per Second	Bandwidth Per Conversation
G.711	20 ms	160	50	80 kbps
G.711	30 ms	240	33	74 kbps
G.729A	20 ms	20	50	24 kbps
G.729A	30 ms	30	33	19 kbps

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

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

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

Table 2-2 shows more accurate bandwidth-provisioning guidelines for voice because it includes Layer 2 overhead.

Table 2-2 *Voice Bandwidth (Including Layer 2 Overhead)*

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

Call-Signaling Traffic

The following list summarizes the key QoS requirements and recommendations for Call-Signaling traffic:

- Call-Signaling traffic should be marked as DSCP CS3 per the QoS Baseline (during migration, it also can be marked the legacy value of DSCP AF31).

- 150 bps (plus Layer 2 overhead) per phone of guaranteed bandwidth is required for voice control traffic; more may be required, depending on the Call-Signaling protocol(s) in use.

Originally, Cisco IP Telephony equipment marked Call-Signaling traffic to DSCP AF31. However, the assured forwarding classes, as defined in RFC 2597, were intended for flows that could be subject to markdown and aggressive dropping of marked-down values. Marking down and aggressively dropping Call-Signaling could result in noticeable delay to dial tone (DDT) and lengthy call-setup times, both of which generally translate into poor user experiences.

Therefore, the QoS Baseline changed the marking recommendation for Call-Signaling traffic to DSCP CS3 because Class-Selector code points, defined in RFC 2474, are not subject to such markdown and aggressive dropping as Assured Forwarding Per-Hop Behaviors are.

Some Cisco IP Telephony products already have begun transitioning to DSCP CS3 for Call-Signaling marking. In this interim period, both code points (CS3 and AF31) should be reserved for Call-Signaling marking until the transition is complete.

Most Cisco IP Telephony products use the Skinny Call-Control Protocol (SCCP) for Call-Signaling. Skinny is a relatively lightweight protocol and, as such, requires only a minimal amount of bandwidth protection (most of the Cisco large-scale lab testing was done by

provisioning only 2 percent for Call-Signaling traffic over WAN and VPN links). However, newer versions of CallManager and SCCP have shown some "bloating" in this signaling protocol, so design recommendations have been adjusted to match (most examples in the design chapters that follow have been adjusted to allocate 5 percent for Call-Signaling traffic). This is a normal part of QoS evolution: As applications and protocols continue to evolve, so do the QoS designs required to accommodate them.

Other Call-Signaling protocols include (but are not limited to) H.225 and H.245, the Session Initiated Protocol (SIP), and the Media Gateway Control Protocol (MGCP). Each Call-Signaling protocol has unique TCP and UDP ports and traffic patterns that should be taken into account when provisioning QoS policies for them.

QoS Requirements of Video

Two main types of video traffic exist: Interactive-Video (videoconferencing) and Streaming-Video (both unicast and multicast). Each type of video is examined separately.

Interactive-Video

When provisioning for Interactive-Video (video conferencing) traffic, the following guidelines are recommended:

- Interactive-Video traffic should be marked to DSCP AF41; excess videoconferencing traffic can be marked down by a policer to AF42 or AF43.
- Loss should be no more than 1 percent.
- One-way latency should be no more than 150 ms.
- Jitter should be no more than 30 ms.
- Assign Interactive-Video to either a preferential queue or a second priority queue (when supported); when using Cisco IOS LLQ, overprovision the minimum-priority bandwidth guarantee to the size of the videoconferencing session plus 20 percent. (For example, a 384-kbps videoconferencing session requires 460 kbps of guaranteed priority bandwidth.)

Because IP videoconferencing (IP/VC) includes a G.711 audio codec for voice, it has the same loss, delay, and delay-variation requirements as voice—but the traffic patterns of videoconferencing are radically different from those of voice.

For example, videoconferencing traffic has varying packet sizes and extremely variable packet rates. These are illustrated in Figures 2-2 and 2-3.

Figure 2-2 *Videoconferencing Traffic Packet-Size Breakdown*
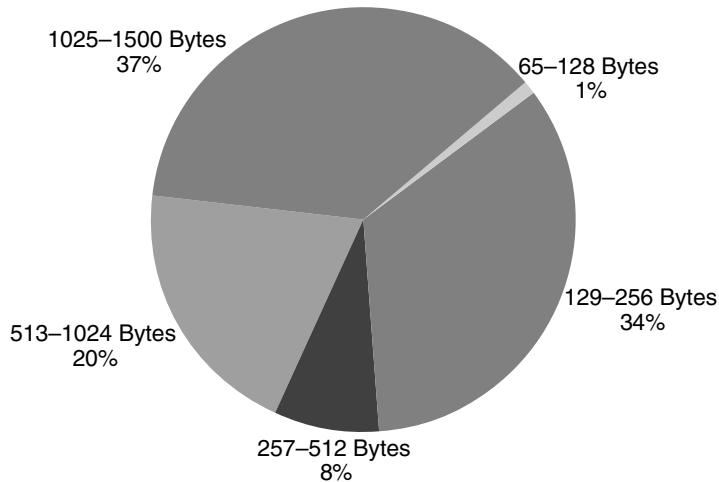
Figure 2-3 *Videoconferencing Traffic Rates (384-kbps Session Example)*
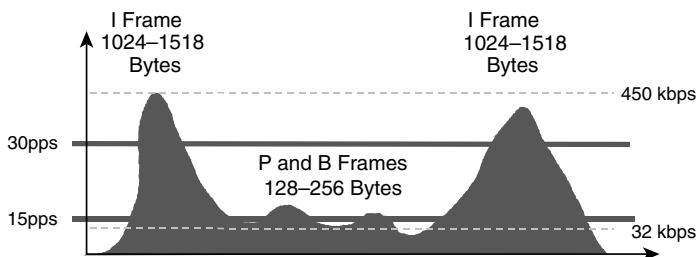
The videoconferencing rate is the sampling rate of the video stream, not the actual bandwidth that the video call requires. In other words, the data payload of videoconferencing packets is filled with 384 kbps of voice plus video samples. IP, UDP, and RTP headers (40 bytes per packet, uncompressed) need to be included in IP/VC bandwidth provisioning, as does the Layer 2 overhead of the media in use. Because (unlike VoIP) IP/VC packet sizes and rates vary, the header overhead percentage also varies, so an absolute value of overhead

cannot be calculated accurately for all streams. However, testing has shown that a conservative rule of thumb for IP/VC bandwidth provisioning is to assign an LLQ bandwidth equivalent to the IP/VC rate plus 20 percent. For example, a 384-kbps IP/VC stream adequately is provisioned with an LLQ of 460 kbps.

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

Streaming-Video

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

- Streaming-Video (whether unicast or multicast) should be marked to DSCP CS4, as designated by the QoS Baseline.

- Loss should be no more than 5 percent.

- Latency should be no more than 4 to 5 seconds (depending on the video application's buffering capabilities).

- There are no significant jitter requirements.

- Guaranteed bandwidth (CBWFQ) requirements depend on the encoding format and rate of the video stream.

- Streaming-Video is typically unidirectional; therefore, remote branch routers might not require provisioning for Streaming-Video traffic on their WAN or VPN edges (in the direction of branch to campus).

- Nonorganizational Streaming-Video applications (either unicast or multicast), such as entertainment video content, may be marked as Scavenger—DSCP CS1, provisioned in the Scavenger traffic class and assigned a minimal bandwidth (CBWFQ) percentage. For more information, see the "Scavenger Class" section, later in this chapter.

Streaming-Video applications have more lenient QoS requirements because they are not delay sensitive (the video can take several seconds to cue up) and are largely not jitter sensitive (because of application buffering). However, Streaming-Video might contain valuable content, such as e-learning applications or multicast company meetings, in which case it requires service guarantees.

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

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

Nonorganizational video content (or video that's strictly entertainment oriented in nature, such as movies, music videos, humorous commercials, and so on) might be considered for Scavenger service, meaning that these streams will play if bandwidth exists, but they will be the first to go during periods of congestion.

QoS Requirements of Data

Hundreds of thousands of data applications exist on the Internet, in all shapes and sizes. Some are TCP, others are UDP; some are delay sensitive, others are not; some are bursty in nature, others are steady; some are lightweight, others are bandwidth hogs—the list goes on.

Data traffic characteristics vary from one application to another, as illustrated in Figure 2-4, which compares an enterprise resource planning (ERP) application (Oracle) with another (SAP).
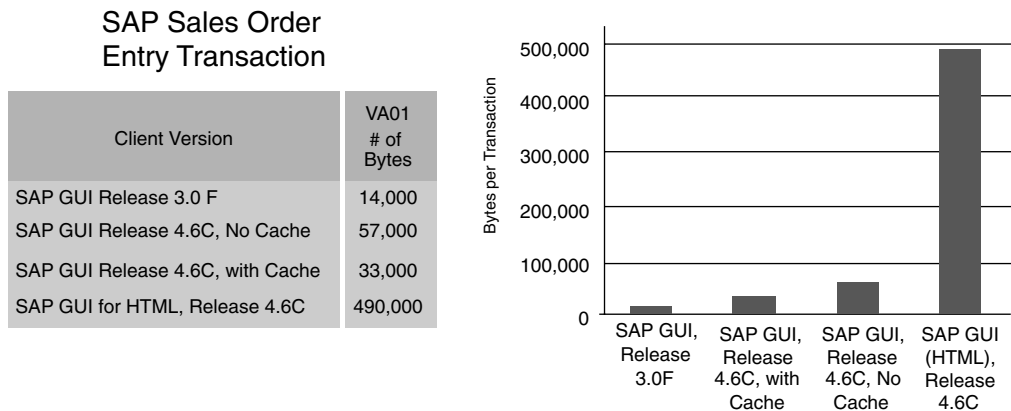
Figure 2-4 *Data Application Differences*
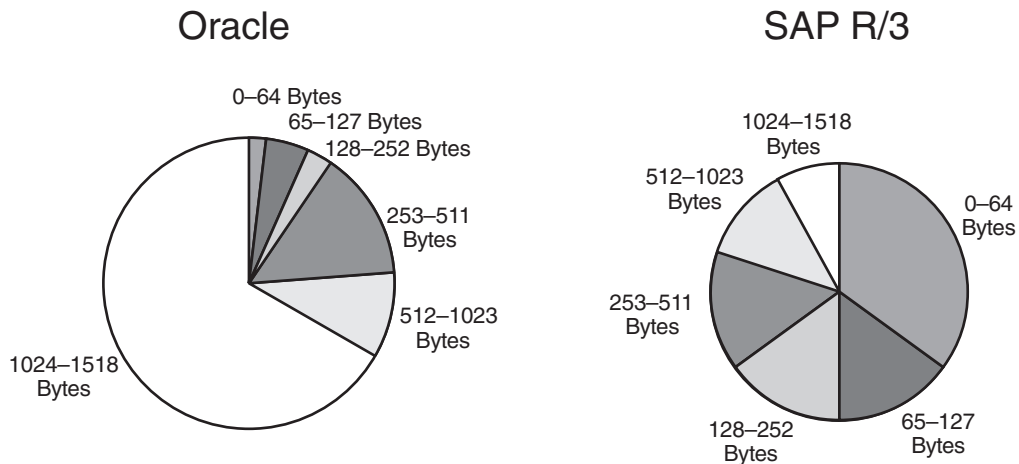
To make the matter even more complicated, it is crucial to recognize that, just as applications vary one from another, even the same application can vary significantly from one *version* to another.

A brief anecdote speaks to this point: After a southern California semiconductor company provisioned QoS for Voice and Mission-Critical Data (SAP R/3), everything went well for about six months. At that point, users began complaining of excessive delays in completing basic transactions. Operations that previously required a second or less to complete were taking significantly longer. The application teams blamed the networking teams, claiming that "QoS was broken." Further investigation produced the information in the following graph, shown in Figure 2-5.

Figure 2-5   *Data Application Version Differences*

SAP Sales Order
Entry Transaction

Client Version	VA01 # of Bytes
SAP GUI Release 3.0 F	14,000
SAP GUI Release 4.6C, No Cache	57,000
SAP GUI Release 4.6C, with Cache	33,000
SAP GUI for HTML, Release 4.6C	490,000

The Mission-Critical Data application—in this instance, SAP—had been upgraded from version 3.0F to 4.6C. As a result, a basic order-entry transaction required 35 times more traffic than the original version. Additional provisioning and policy tuning was required to accommodate the new version of the same application.

Given this reality, the question on how best to provision QoS for data is a daunting one. After wrestling with this question for several years, the authors of the QoS Baseline came up with four main classes of data traffic, according to their general networking characteristics and requirements. These classes are Best-Effort, Bulk Data, Transactional Data/Interactive Data and (Locally-Defined) Mission-Critical Data. Each of these classes is examined in more detail in the following sections.

Best-Effort Data

When addressing the QoS needs of Best-Effort traffic, the following guidelines are recommended:

- Best-Effort traffic should be marked to DSCP 0.

- Adequate bandwidth should be assigned to the Best-Effort class as a whole because the majority of applications default to this class. It is recommended to reserve at least 25 percent for Best-Effort traffic.

The Best-Effort class is the default class for all data traffic. Only if an application has been selected for preferential or deferential treatment is it removed from the default class.

In 2003, one Wall Street financial company did an extensive study to identify and categorize the number of different applications on its networks. It found more than 3000 discrete applications traversing its infrastructure. Further research has shown that this is not uncommon for larger enterprises. Therefore, because enterprises have several hundred—if not thousands of—data applications running over their networks (of which the majority default to the Best-Effort class), adequate bandwidth needs to be provisioned for this default class to handle the sheer volume of applications that are included in it. Otherwise, applications that default to this class easily are drowned out, typically resulting in an increased number of calls to the networking help desk from frustrated users. It is therefore recommended that at least 25 percent of a link's bandwidth be reserved for the default Best-Effort class.

Bulk Data

When addressing the QoS needs of Bulk Data traffic, the following guidelines are recommended:

- Bulk Data traffic should be marked to DSCP AF11; excess Bulk Data traffic can be marked down by a policer to AF12 or AF13.

- Bulk Data traffic should have a moderate bandwidth guarantee but should be constrained from dominating a link.

The Bulk Data class is intended for applications that are relatively noninteractive and not drop sensitive, and that typically span their operations over a long period of time as background occurrences. Such applications include FTP, e-mail, backup operations, database synchronizing or replicating operations, video content distribution, and any other type of application in which users typically cannot proceed because they are waiting for the completion of the operation (in other words, a background operation).

The advantage of provisioning moderate bandwidth guarantees to Bulk Data applications (instead of applying policers to them) is that Bulk Data applications dynamically can take advantage of unused bandwidth and thus can speed up their operations during nonpeak periods. This, in turn, reduces the likelihood that they will bleed into busy periods and absorb inordinate amounts of bandwidth for their non-time-sensitive operations.

Transactional Data/Interactive Data

When addressing the QoS needs of Transactional Data and Interactive Data traffic, the following guidelines are recommended:

- Transactional Data traffic should be marked to DSCP AF21; excess Transactional Data traffic can be marked down by a policer to AF22 or AF23.

- Transactional Data traffic should have an adequate bandwidth guarantee for the interactive, foreground operations that it supports.

The Transactional Data/Interactive Data class is a combination of two similar types of applications: Transactional Data client/server applications and interactive messaging applications. For the sake of simplicity, this class is referred to as Transactional Data only.

The response-time requirement separates Transactional Data client/server applications from generic client/server applications. For example, with Transactional Data client/server applications (such as SAP, PeopleSoft, and Oracle), the user waits for the operation to complete before proceeding (in other words, the transaction is a foreground operation). E-mail is not considered a Transactional Data client/server application because most e-mail operations happen in the background, and users usually do not notice even delays of several hundred milliseconds in mailspool operations.

Locally Defined Mission-Critical Data

When addressing the QoS needs of Locally-Defined Mission-Critical Data traffic, the following guidelines are recommended:

- Locally-Defined Mission-Critical Data traffic should be marked to DSCP AF31; excess Mission-Critical Data traffic can be marked down by a policer to AF32 or AF33. However, Cisco IP Telephony equipment currently is using DSCP AF31 to mark Call-Signaling traffic; until all Cisco IPT products mark Call-Signaling to DSCP CS3, a temporary placeholder code point, DSCP 25, can be used to identify Locally-Defined Mission-Critical Data traffic.

- Locally-Defined Mission-Critical Data traffic should have an adequate bandwidth guarantee for the interactive, foreground operations that it supports.

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

For example, an enterprise might have properly provisioned Oracle, SAP, BEA, and Siebel within its Transactional Data class. However, the majority of its revenue might come from SAP, so it might want to give this Transactional Data application an even higher level of

preference by assigning it to a dedicated class (such as the Locally-Defined Mission-Critical class).

Because the admission criteria for this class is nontechnical (being determined by business relevance and organizational objectives), the decision about which application(s) should be assigned to this special class easily can become an organizationally and politically charged debate. It is recommended to assign as few applications to this class (from the Transactional Data class) as possible. In addition, it is recommended that executive endorsement for application assignments to the Locally-Defined Mission-Critical class be obtained: The potential for QoS deployment derailment exists without such an endorsement.

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

Based on these definitions, Table 2-3 shows some applications and their generic networking characteristics, which determine what data application class they are best suited to.

Table 2-3   *Data Applications by Class*

Application Class	Example Applications	Application/Traffic Properties	Packet/Message Sizes
Interactive PHB: AF2	Telnet, Citrix, Oracle Thin-Clients, AOL Instant Messenger, Yahoo! Instant Messenger, PlaceWare (Conference), Netmeeting Whiteboard.	Highly interactive applications with tight user-feedback requirements.	Average message size < 100 bytes. Max message size < 1 KB.
Transactional PHB: AF2	SAP, PeopleSoft—Vantive, Oracle—Financials, Internet Procurement, B2B, Supply Chain Management, Application Server, Oracle 8i Database, Ariba Buyer, I2, Siebel, E.piphany, Broadvision, IBM Bus 2 Bus, Microsoft SQL, BEA Systems, DLSw+.	Transactional applications typically use a client/server protocol model. User-initiated, client-based queries are followed by server response. The query response can consist of many messages between client and server. The query response can consist of many TCP and FTP sessions running simultaneously (for example, HTTP-based applications).	Depends on application; could be anywhere from 1 KB to 50 MB.

Table 2-3   *Data Applications by Class (Continued)*

Application Class	Example Applications	Application/Traffic Properties	Packet/Message Sizes
Bulk PHB: AF1	Database syncs, network-based backups, Lotus Notes, Microsoft Outlook, e-mail download (SMTP, POP3, IMAP, Exchange), video content distribution, large FTP file transfers.	Long file transfers. Always invokes TCP congestion management.	Average message size 64 KB or greater.
Best-Effort PHB: Default	All noncritical traffic, HTTP web browsing, other miscellaneous traffic.		

DLSw+ Considerations

Some enterprises support legacy IBM equipment that requires data-link switching plus (DLSw+) to operate across an enterprise environment.

In such cases, it is important to recognize that DLSw+ traffic, by default, is marked to IP Precedence 5 (DSCP CS5). This default marking could interfere with VoIP provisioning because both DSCP EF and DSCP CS5 share the same IP Precedence, 802.1Q/p CoS, and MPLS EXP value (of 5). Therefore, it is recommended to re-mark DLSw+ traffic away from this default value of IP Precedence 5.

Unfortunately, at the time of writing, Cisco IOS does not support marking DSCP values within the DLSw+ peering statements; it supports only the marking of type of service (ToS) values using the **dlsw tos map** command.

NOTE   To explain this behavior from a historical perspective, when DLSw+ was developed, *ToS* was a term loosely used to refer to the first 3 bits (the IP Precedence bits) of the IP ToS byte, as defined in RFC 791. For many years, these were typically the only bits of the ToS byte in use (the others almost always were set to 0), so it seemed to make sense at the time.

However, with the development of newer standards defining the use of the first 6 bits of the IP ToS byte for DiffServ markings (RFC 2474) and the last 2 bits for IP explicit congestion notification (RFC 3168), using the term *ToS* to refer to only the first 3 bits (the IP Precedence bits) of the IP ToS byte has become increasingly inaccurate.

However, marking DLSw+ to an IP Precedence/class selector value could interfere with other QoS Baseline recommended markings. For example, if DLSw+ is marked to IPP 1, it would be treated as Scavenger traffic; if it is marked to IPP 2, it could interfere with Network-Management traffic; if it is marked to IPP 3, it could interfere with Call-Signaling; if it is marked to IPP 4, it would be treated as Streaming-Video traffic; and if it is marked to IPP 6 or 7, it could interfere with routing or Network Control traffic.

Therefore, a two-step workaround is recommended for marking DLSw+ traffic:

Step 1   Disable native DLSw+ ToS markings with the **dlsw tos disable** command.

Step 2   Identify DLSw+ traffic either with access lists (matching the DLSw+ TCP ports 1981 to 1983 or 2065) or with the **match protocol dlsw** command.

When DLSw+ traffic is identified, it can be marked as either Transactional (AF21) or Mission-Critical Data (DSCP 25), depending on the organization's preference.

QoS Requirements of the Control Plane

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

IP Routing

When addressing the QoS needs of IP routing traffic, the following guidelines are recommended:

- IP routing traffic should be marked to DSCP CS6; this is default behavior on Cisco IOS platforms.
- Interior gateway protocols usually adequately are protected with the Cisco IOS internal PAK_priority mechanism. Exterior gateway protocols, such as BGP, are recommended to have an explicit class for IP routing with a minimal bandwidth guarantee.

Cisco IOS automatically marks IP routing traffic to DSCP CS6.

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

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

It is important to note that although exterior gateway protocol (EGP) traffic, such as Border Gateway Protocol (BGP) traffic, is marked by default to DSCP CS6, it does not receive such PAK_priority preferential treatment and might need to be protected explicitly to maintain peering sessions.

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

Network-Management

When addressing the QoS needs of Network-Management traffic, the following guidelines are recommended:

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

Network-Management traffic is important in performing trend and capacity analyses and troubleshooting. Therefore, a separate minimal bandwidth queue can be provisioned for Network-Management traffic, which could include SNMP, NTP, Syslog, and NFS and other management applications.

Scavenger Class

When addressing the QoS treatment of Scavenger traffic, the following guidelines are recommended:

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

The Scavenger class is intended to provide deferential services, or less-than best-effort services, to certain applications. Applications assigned to this class have little or no contribution to the organizational objectives of the enterprise and are typically entertainment oriented in nature. These include peer-to-peer media-sharing applications (KaZaa, Morpheus,

Groekster, Napster, iMesh, and so on), gaming applications (Doom, Quake, Unreal Tournament, and so on), and any entertainment video applications.

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

The Scavenger class is a critical component to the DoS and worm mitigation strategy, discussed next.

DoS and Worm Mitigation Strategy Through Scavenger Class QoS

Worms are nothing new; they have been around in some form since the beginning of the Internet and steadily have been increasing in complexity, as shown in Figure 2-6.

Figure 2-6 *Business Security Threat Evolution*
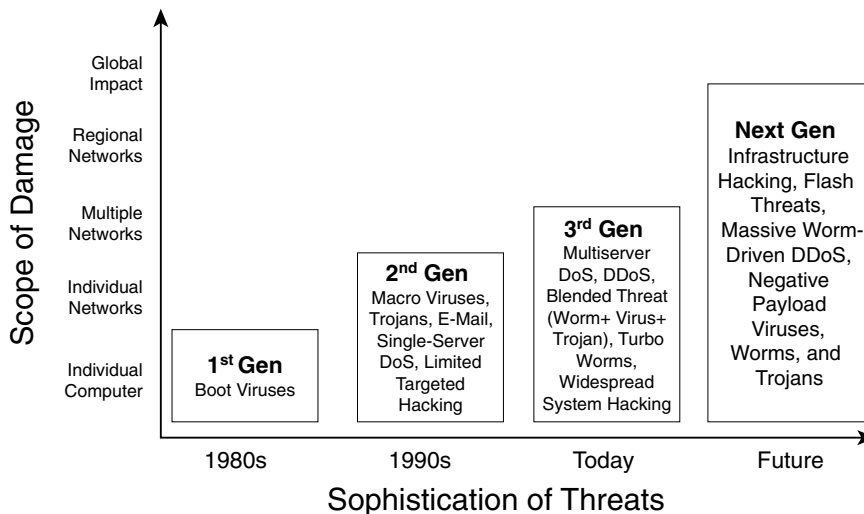
Particularly since 2002, there has been an exponential increase not only in the frequency of DoS and worm attacks, but also in their relative sophistication and scope of damage. For example, more than 994 new Win32 viruses and worms were documented in the first half of 2003, more than double the 445 documented in the first half of 2002. Some of these more recent worms are shown in Figure 2-7.

Figure 2-7 *Recent Internet Worms*
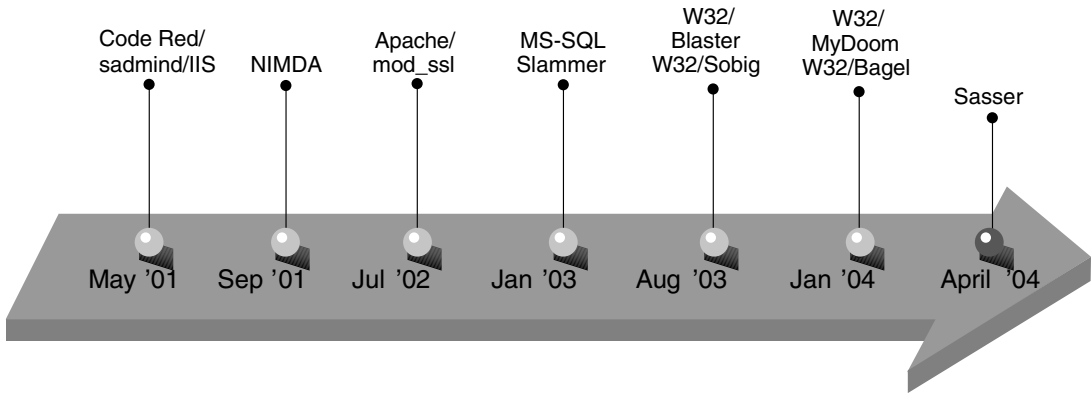
DoS or worm attacks can be categorized into two main classes:

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

Spoofing attacks best are addressed by authentication and encryption technologies; flooding attacks, on the other hand, can be mitigated using QoS technologies.

The majority of flooding attacks target PCs and servers, which, when infected, target other PCs and servers, thus multiplying traffic flows. Network devices themselves are not usually the direct targets of attacks. But the rapidly multiplying volumes of traffic flows eventually drown the CPU and hardware resources of routers and switches in their paths, causing denial of service to legitimate traffic flows. The end result is that network devices become indirect victims of the attack. This is illustrated in Figure 2-8.

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

- 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 might be in the path of the worm. This task might be hampered by the attack itself because network devices might become unreachable for administration during the attacks.

Figure 2-8 *Impact of an Internet Worm—Direct and Collateral Damage*
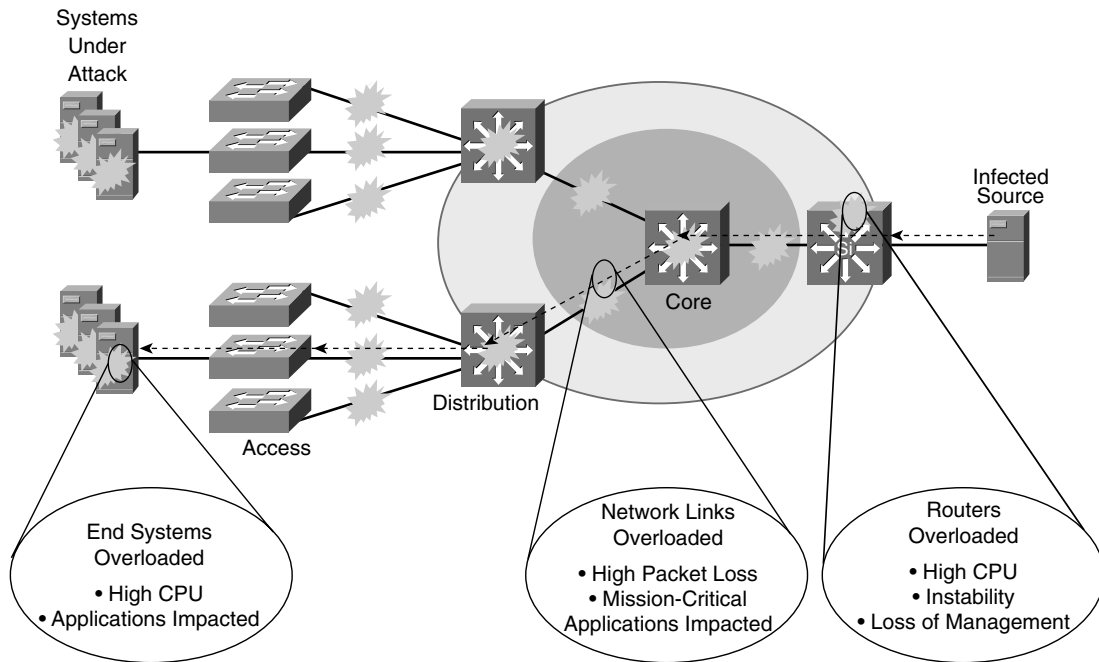
These time lags might not seem long in absolute terms, such as in minutes, but the relative window of opportunity for damage is huge. For example, in 2003, the number of hosts infected with the Slammer worm (a Sapphire worm variant) doubled every 8.5 seconds on average, infecting more than 75,000 hosts in just 11 minutes and performing scans of 55 million more hosts within the same time period.

NOTE Interestingly, a 2002 CSI/FBI report stated that the majority of network attacks occur from within an organization, typically by disgruntled employees.

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

In this respect, the policers would be fairly "dumb." They would not be matching specific network characteristics of specific types of attacks, but they simply would be metering traffic volumes and responding to abnormally high volumes as close to the source as possible. The simplicity of this approach negates the need for the policers to be programmed with knowledge of the specific details of how the attack is being generated or propagated. It is precisely this "dumbness" of such Access-Layer policers that allows them to maintain relevancy as worms mutate and become more complex: The policers don't care how the traffic was generated or what it looks like; all they care about is how much traffic is being put onto the wire. Therefore, they continue to police even advanced worms that continually change the tactics of how traffic is being generated.

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

NOTE	It is important to recognize that this value (≤ 5 percent) for normal access-edge utilization by endpoints is just an example value. This value would likely vary from industry vertical to vertical, and from enterprise to enterprise.

It is important to recognize that what is being proposed is not to police all traffic to 5 Mbps and automatically drop the excess. If that were the case, there would not be much reason for deploying Fast Ethernet or Gigabit Ethernet switch ports to endpoint devices because even 10BASE-T Ethernet switch ports would have more uplink capacity than a 5 Mbps policer-enforced limit. Furthermore, such an approach supremely would penalize legitimate traffic that exceeded 5 Mbps on a Fast Ethernet switch port.

A less draconian approach is to couple Access-Layer policers with hardware and software (campus, WAN, and VPN) queuing policies, with both sets of policies provisioning for a less-than best-effort Scavenger class.

This would work by having Access-Layer policers mark down out-of-profile traffic to DSCP CS1 (Scavenger) and then have all congestion-management policies (whether in Catalyst hardware or in Cisco IOS Software) provision a less-than best-effort service for any traffic marked to DSCP CS1.

Let's examine how this might work, for both legitimate traffic exceeding the Access-Layer policer's watermark and illegitimate excess traffic (the result of a DoS or worm attack).

In the former case, imagine that the PC generates more than 5 Mbps of traffic, perhaps because of a large file transfer or backup. Because there is generally abundant capacity within the campus to carry the traffic, congestion (under normal operating conditions) is

rarely, if ever, experienced. Typically, the uplinks to the distribution and core layers of the campus network are Gigabit Ethernet, which requires 1000 Mbps of traffic from the Access-Layer switch to create congestion. If the traffic was destined to the far side of a WAN or VPN link (which are rarely more than 5 Mbps in speed), dropping would occur even without the Access-Layer policer, simply because of the campus/WAN speed mis-match and resulting bottleneck. TCP's sliding-windows mechanism eventually would find an optimal speed (less than 5 Mbps) for the file transfer.

To make a long story short, Access-Layer policers that mark down out-of-profile traffic to Scavenger (CS1) would not affect legitimate traffic, aside from the obvious re-marking. No reordering or dropping would occur on such flows as a result of these policers (that would not have occurred anyway).

In the latter case, the effect of Access-Layer policers on traffic caused by DoS or worm attacks is quite different. As hosts become infected and traffic volumes multiply, congestion might be experienced even within the campus. If just 11 end-user PCs on a single switch begin spawning worm flows to their maximum Fast Ethernet link capacities, the GE uplink from the Access-Layer switch to the Distribution-Layer switch will congest and queuing or reordering will engage. At such a point, VoIP and critical data applications, and even Best-Effort applications, would gain priority over worm-generated traffic (and Scavenger traffic would be dropped the most aggressively); network devices would remain accessible for administration of patches, plugs, and ACLs required to fully neutralize the specific attack.

WAN links also would be protected: VoIP, critical data, and even best-effort flows would continue to receive priority over any traffic marked down to Scavenger/CS1. This is a huge advantage because generally WAN links are the first to be overwhelmed by DoS and worm attacks. The bottom line is that Access-Layer policers significantly mitigate network traffic generated by DoS or worm attacks.

It is important to recognize the distinction between mitigating an attack and preventing it entirely: The strategy being presented does not guarantee that no denial of service or worm attacks ever will happen, but it can reduce the risk and impact that such attacks could have on the network infrastructure.

Principles of QoS Design

The richness of the Cisco QoS toolset allows for a myriad of QoS design and deployment options. However, a few succinct design principles can help simplify strategic QoS designs and lead to an expedited, cohesive, and holistic end-to-end deployment. Some of these design principles are summarized here; others, which are LAN-, WAN- or VPN-specific, are covered in detail in their respective design chapters.

General QoS Design Principles

A good place to begin is to decide which comes first: the cart or the horse. The horse, in this context, serves to pull the cart and is the enabler for this objective. Similarly, QoS technologies are simply the enablers to organizational objectives. Therefore, the way to begin a QoS deployment is not by glossing over the QoS toolset and picking à la carte tools to deploy. In other words, do not enable QoS features simply because they exist. Instead, start from a high level and clearly define the organizational objectives.

Some questions for high-level consideration include the following:

- Is the objective to enable VoIP only?
- Is video also required? If so, what type(s) of video: interactive or streaming?
- Are some applications considered mission critical? If so, what are they?
- Does the organization want to squelch certain types of traffic? If so, what are they?

All traffic classes specified in the QoS Baseline model except one—the Locally-Defined, Mission-Critical Data application class—are determined by objective networking characteristics. These applications, a subset of the Transactional Data class, are selected for a dedicated, preferential class of service because of their significant impact on the organization's main business objectives.

This is usually a highly subjective evaluation that can excite considerable controversy and dispute. An important principle to remember when assigning applications to the Mission-Critical Data class is that as few applications as possible should be assigned to the Locally-Defined Mission-Critical class.

If too many applications are assigned to it, the Mission-Critical Data class will dampen, and possibly even negate, the value of having a separate class (from Transactional Data). For example, if 10 applications are assigned as Transactional Data (because of their interactive, foreground networking characteristics) and all 10 are determined to be classified as Mission-Critical Data, the whole point of a separate class for these applications becomes moot. However, if only one or two of the Transactional Data applications are assigned to the Mission-Critical Data class, the class will prove highly effective.

Related to this point, it is recommended always to seek executive endorsement of the QoS objectives before design and deployment. By its very nature, QoS is a system of managed unfairness and, as such, almost always creates political and organizational repercussions when implemented. To minimize the effects of such nontechnical obstacles to deployment, which could prevent the QoS implementation altogether, it is recommended to address these political and organizational issues as early as possible and to solicit executive endorsement whenever possible.

As stated previously, it is not mandated that enterprises deploy all 11 classes of the QoS Baseline model; this model is designed to be a forward-looking guide for consideration of the many classes of traffic that have unique QoS requirements. Being aware of this model

can help bring about a smooth expansion of QoS policies to support additional applications as future requirements arise. However, at the time of QoS deployment, the organization needs to clearly define how many classes of traffic are required to meet the organizational objectives.

This consideration should be tempered with the consideration of how many classes of applications the networking administration team feels comfortable with deploying and supporting. Platform-specific constraints or service-provider constraints also might come into play when arriving at the number of classes of service. At this point, it also would be good to consider a migration strategy to allow the number of classes to be expanded smoothly as future needs arise, as illustrated in Figure 2-9.

Figure 2-9   *Example Strategy for Expanding the Number of Classes of Service over Time*
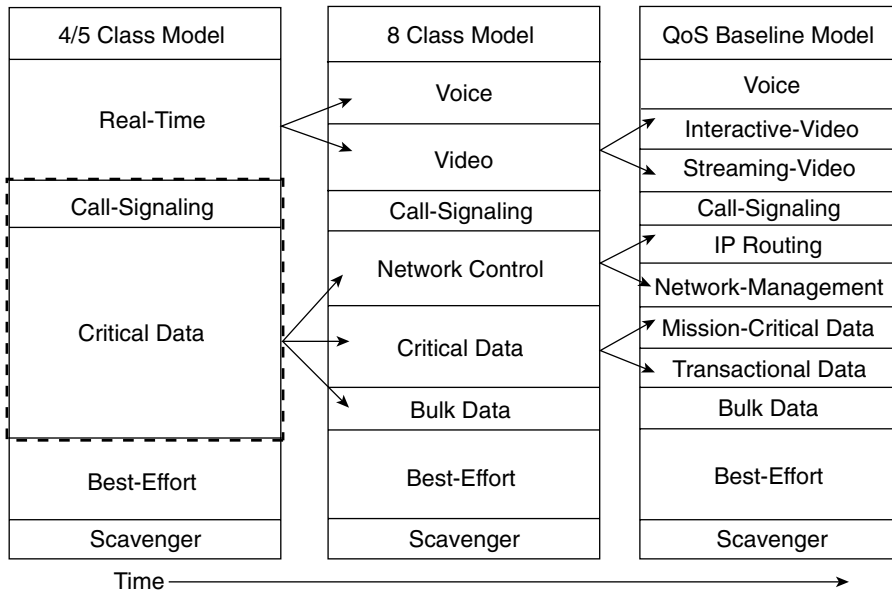
When the number of classes of service has been determined, the details of the required marking, policing, and queuing policies can be addressed. When deciding where to enable such policies, keep in mind that QoS policies always should be performed in hardware instead of software whenever a choice exists.

Cisco IOS routers perform QoS in software, which places incremental loads on the CPU (depending on the complexity and functionality of the policy). Cisco Catalyst switches, on the other hand, perform QoS in dedicated hardware ASICS and, as such, do not tax their main CPUs to administer QoS policies. This allows complex policies to be applied at line rates at even 1-Gbps or 10-Gigabit speeds.

Classification and Marking Principles

When it comes to classifying and marking traffic, an unofficial Differentiated Services design principle is to classify and mark applications as close to their sources as technically and administratively feasible. This principle promotes end-to-end Differentiated Services and per-hop behaviors (PHBs). Sometimes endpoints can be trusted to set CoS and DSCP markings correctly, but, in most cases, it is not a good idea to trust markings that users can set on their PCs (or other similar devices). This is because users easily could abuse provisioned QoS policies if permitted to mark their own traffic. For example, if DSCP EF receives priority services throughout the enterprise, a user easily could configure the PC to mark all traffic to DSCP EF right on the NIC, thus hijacking network-priority queues to service that user's non-real-time traffic. Such abuse easily could ruin the service quality of real-time applications (such as VoIP) throughout the enterprise. For this reason, the clause "as close as . . . *administratively* feasible" is included in the design principle.

Following this rule, it further is 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 at a LAN-to-WAN or VPN edge). An additional constraint to Layer 2 marking is that there is less marking granularity; for example, 802.1Q/p CoS supports only 3 bits (values 0 through 7), as does MPLS EXP. Therefore, only (up to) eight classes of traffic can be supported at Layer 2, and interclass relative priority (such as RFC 2597 assured-forwarding class 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.

Because the line between enterprises and service providers is blurring and the need for interoperability and complementary QoS markings is critical, it is recommended to follow standards-based DSCP PHB markings to ensure interoperability and future expansion. The QoS Baseline marking recommendations are standards based, making it easier for enterprises adopting these markings to interface with service provider classes of service. Network mergers are also easier to manage when standards-based DSCP markings are used, whether these mergers are the result of acquisitions, partnerships, or strategic alliances.

Policing and Markdown Principles

There is little sense in forwarding unwanted traffic only to police and drop it at a subsequent node. This is especially the case when the unwanted traffic is the result of DoS or worm attacks. The overwhelming volumes of traffic that such attacks can create readily can drive network device processors to their maximum levels, causing network outages. Therefore, it is recommended to police traffic flows as close to their sources as possibl*e*. This principle applies to legitimate flows also because DoS and worm-generated traffic might be masquerading under legitimate, well-known TCP and UDP ports, causing extreme amounts of

traffic to be poured onto the network infrastructure. Such excesses should be monitored at the source and marked down appropriately.

Whenever supported, markdown should be done according to standards-based rules, such as RFC 2597 ("Assured Forwarding PHB Group"). In other words, whenever supported, traffic marked to AFx1 should be marked down to AFx2 or AFx3. For example, in the case of a single-rate policer, excess traffic originally marked AF11 should be marked down to AF12. In the case of a dual-rate policer (as defined in RFC 2698), excess traffic originally marked AF11 should be marked down to AF12, and violating traffic should be marked down further to AF13. Following such markdowns, congestion-management policies, such as DSCP-based WRED, should be configured to drop AFx3 more aggressively than AFx2, which, in turn, is dropped more aggressively than AFx1.

However, at the time of writing, Cisco Catalyst switches do not perform DSCP-based WRED, so this standards-based strategy cannot be implemented fully. As an alternative workaround, single-rate policers can be configured to mark down excess traffic to DSCP CS1 (Scavenger); dual-rate policers can be configured to mark down excess traffic to AFx2, while marking down violating traffic to DSCP CS1. Such workarounds yield an overall similar effect as the standards-based policing model. However, when DSCP-based WRED is supported on all routing and switching platforms, it would be more standards compliant to mark down assured-forwarding classes by RFC 2597 rules.

Queuing and Dropping Principles

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

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

Another class of traffic that requires special consideration when provisioning queuing is the Real-Time or Strict-Priority class (which corresponds to RFC 3246, "An Expedited Forwarding Per-Hop Behavior"). The amount of bandwidth assigned to the Real-Time

queuing class is variable. However, if too much traffic is assigned for strict-priority queuing, the overall effect is a dampening of QoS functionality for non-real-time applications.

The goal of convergence cannot be overemphasized: to enable voice, video, and data to coexist transparently on a single network. When real-time applications (such as Voice or Interactive-Video) dominate a link (especially a WAN/VPN link), data applications will fluctuate significantly in their response times, destroying the transparency of the "converged" network.

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

Cisco IOS Software allows the abstraction (and, thus, configuration) of multiple (strict-priority) low-latency queues. In such a multiple-LLQ context, this design principle applies to the sum of all LLQs: They should be within one-third of a link's capacity.

| NOTE | This strict-priority queuing rule (limit to 33 percent) is simply a best-practice design recommendation; it is not a mandate. In some cases, specific business objectives cannot be met while holding to this recommendation. In such cases, enterprises must provision according to their detailed requirements and constraints. However, it is important to recognize the trade-offs involved with overprovisioning strict-priority traffic with respect to the negative performance impact on response times in non-real-time applications. |

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

The Real-Time, Best-Effort, and Scavenger classes queuing best-practice principles are illustrated in Figure 12-10.

Figure 2-10 *Real-Time, Best-Effort, and Scavenger Queuing Rules*
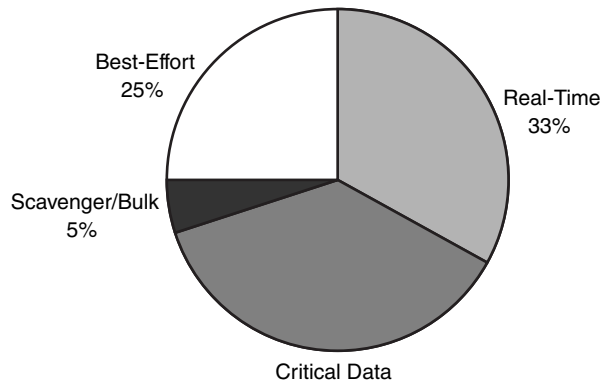
Some platforms support different queuing structures than others. To ensure consistent PHBs, configure consistent queuing policies according to platform capabilities.

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

- Real-Time (≤ 33 percent)
- Critical Data
- Best-Effort (≥ 25 percent)
- Scavenger/Bulk(< 5 percent)

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

- Voice (≤ 18 percent)
- Interactive-Video (≤ 15 percent)
- Internetwork Control
- Call-Signaling
- Mission-Critical Data
- Transactional Data
- Network-Management
- Streaming-Video Control

- Best-Effort (≥ 25 percent)
- Bulk Data (4 percent)
- Scavenger (1 percent)

Figure 2-11 illustrates the interrelationship between these compatible queuing models.

Figure 2-11 *Compatible 4-Class and 11-Class Queuing Models Following Real-Time, Best-Effort, and Scavenger Class Queuing Rules*
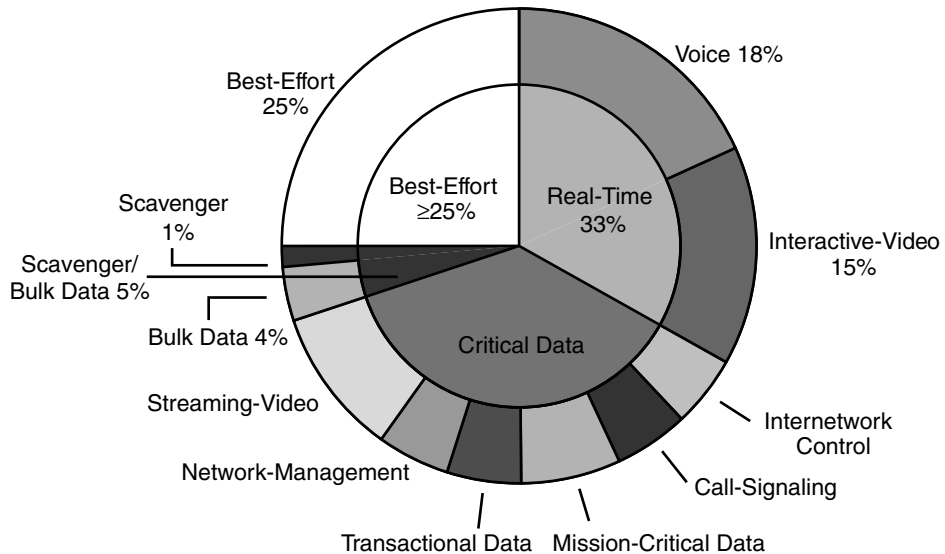
In this manner, traffic will receive compatible queuing at each node, regardless of platform capabilities—which is the overall objective of DiffServ per-hop behavior definitions.

Whenever supported, it is recommended to enable WRED (preferably DSCP-based WRED) on all TCP flows. In this manner, WRED congestion avoidance will prevent TCP global synchronization and will increase overall throughput and link efficiency. Enabling WRED on UDP flows is optional.

DoS and Worm Mitigation Principles

Whenever part of the organization's objectives is to mitigate DoS and worm attacks through Scavenger-class QoS, the following best practices apply.

First, the network administrators need to profile applications to determine what constitutes normal versus abnormal flows, within a 95 percent confidence interval. Thresholds

differentiating normal and abnormal flows vary from enterprise to enterprise and from application to application. Caution must be extended not to overscrutinize traffic behavior because this could be time and resource exhaustive and easily could change from one day to the next. Remember, the presented Scavenger-class strategy will not apply a penalty to legitimate traffic flows that exceed thresholds (aside from re-marking); only sustained, abnormal streams generated simultaneously by multiple hosts (highly indicative of DoS and worm attacks) are subject to aggressive dropping, and only after legitimate traffic has been serviced.

To contain such abnormal flows, it is recommended to deploy campus Access-Edge policers to re-mark abnormal traffic to Scavenger (DSCP CS1). Additionally, whenever Catalyst 6500s with Supervisor 720s are deployed in the distribution layer, it is recommended to deploy a second line of policing defense, at the distribution layer via per-user microflow policing.

To complement these re-marking policies, it is necessary to enforce end-to-end Scavenger-class queuing policies, where flows marked as Scavenger will receive a less-than best-effort service whenever congestion occurs.

It is important to note that even when Scavenger-class QoS has been deployed end to end, this strategy only mitigates DoS and worm attacks and does not prevent them or remove them entirely. Therefore, it is critical to overlay security, firewall, intrusion detection, and identity systems, along with Cisco Guard and Cisco Security Agent solutions, on top of the QoS-enabled network infrastructure.

Deployment Principles

After the QoS designs have been finalized, it is vital that the networking team thoroughly understand the QoS features and syntax before enabling features on production networks. Such knowledge is critical for both deployment and troubleshooting QoS-related issues.

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

When testing has validated the designs, it is recommended to schedule network downtime to deploy QoS features. Although QoS is required end to end, it does not have to be deployed end to end at a single instance. A pilot network segment can be selected for an initial deployment, and, pending observation, the deployment can be expanded in stages to encompass the entire enterprise. A rollback strategy always is recommended, to address unexpected issues that arise from the QoS deployment.

Summary

This chapter began by reviewing the QoS requirements of voice, video, and data applications.

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

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

Control plane requirements, such as provisioning moderate bandwidth guarantees for IP routing protocols and network-management protocols, should not be overlooked.

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

A less-than best-effort Scavenger class of traffic was introduced, and a strategy for using this class for DoS and worm mitigation was presented. Specifically, flows can be monitored at the campus Access-Edge, and out-of-profile flows can be marked down to the Scavenger marking (of DSCP CS1). To complement these policers, queues providing a less-than best-effort Scavenger service during periods of congestion are deployed in the LAN, WAN, and VPN.

The chapter concluded with a set of general best-practice principles relating to QoS planning, classification, marking, policing, queuing, and deployment. These best practices include the following:

- Clearly defining the organization's business objectives to be addressed by QoS

- Selecting an appropriate number of service classes to meet these business objectives

- Soliciting executive endorsement, whenever possible, of the traffic classifications, especially when determining any mission-critical applications

- Performing QoS functions in (Catalyst switch) hardware instead of (Cisco IOS router) software, whenever possible

- Classifying traffic as close to the source as administratively feasible, preferably at Layer 3 with standards-based DSCP markings

- Policing traffic as close to the source as possible, following standards-based rules (such as RFC 2597, "Assured Forwarding Markdown"), whenever possible

- Provisioning at least one-quarter of a link to service Best-Effort traffic

- Provisioning no more than one-third of a link to service real-time and strict-priority traffic

- Provisioning a less-than best-effort Scavenger queue, which should be assigned as low of a bandwidth allocation as possible

- Understanding and thoroughly testing desired QoS features in conjunction with features already enabled on the production network

- Deploying end-to-end QoS in stages during scheduled network downtime, with a recommended rollback strategy

Further Reading

Standards:

- RFC 791, "Internet Protocol Specification": http://www.ietf.org/rfc/rfc791.

- RFC 2474, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers": http://www.ietf.org/rfc/rfc2474.

- RFC 2597, "Assured Forwarding PHB Group": http://www.ietf.org/rfc/rfc2597.

- RFC 2697, "A Single Rate Three Color Marker": http://www.ietf.org/rfc/rfc2697.

- RFC 2698, "A Two Rate Three Color Marker": http://www.ietf.org/rfc/rfc2698.

- RFC 3168, "The Addition of Explicit Congestion Notification (ECN) to IP": http://www.ietf.org/rfc/rfc3168.

- RFC 3246, "An Expedited Forwarding PHB (Per-Hop Behavior)": http://www.ietf.org/rfc/rfc3246.

Cisco documentation:

- Cisco IOS QoS Configuration Guide, Cisco IOS version 12.3: http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123cgcr/qos_vcg.htm.

- Cisco IOS Configuration Guide—Configuring Data Link Switching Plus, Cisco IOS version 12.3: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fibm_c/bcfpart2/bcfdlsw.htm.

- Understanding how routing updates and Layer 2 control packets are queued on an interface with a QoS service policy (PAK_priority): http://www.cisco.com/warp/public/105/rtgupdates.html.