

Service Provider Quality-of-Service Overview

Introduction

This white paper provides an overview for deploying quality of service (QoS) in the service provider network, including information on Cisco® AutoQoS, class of service (CoS) definitions for different traffic groups, and best-practice procedures.

QoS refers to the capability to provide predictable performance and better service to specific classes of network traffic, and is an essential tool for service providers to meet enterprise demands and expectations for voice, data, video, etc. Primary goals of QoS include dynamic bandwidth allocation for mission-critical applications and prioritization of delay-sensitive traffic such as voice and video. QoS mechanisms include queuing, network congestion avoidance, traffic shaping, and packet classification.

Service providers can offer scalable voice and video deployments and advanced Layer 3 QoS capabilities. Proactive monitoring, performance management, project management, customer service resources, installation and support services, and detailed network reports are other benefits that can be provided to enterprise customers and that can be applied to achieve the required QoS for voice over IP (VoIP), videoconferencing, video on demand (VoD), and other quality-sensitive applications and services.

Customer networks exist to service application requirements and end users efficiently. The tremendous growth of the Internet and corporate intranet; the wide variety of new bandwidth-hungry applications; and convergence of data, voice, and video traffic over consolidated IP infrastructures has had a major impact on the ability of networks to provide predictable, measurable, and guaranteed services to these applications. Achieving the required QoS through the proper management of network delays, bandwidth requirements, and packet-loss parameters, while maintaining simplicity, scalability, and manageability of the network is the fundamental solution to running an infrastructure that serves business applications end to end.

QoS refers to the ability of a network to provide better service to selected network traffic over various underlying technologies including Frame Relay, ATM, Ethernet, and 802.1 networks; SONET; and IP-routed networks. It is a collection of technologies that allows applications to request and receive predictable service levels in terms of data throughput capacity (bandwidth), latency variations (jitter), and delay.

QoS ensures that critical business applications perform effectively.



CoS is a priority output queuing mechanism that allows packet prioritization based on protocol, port, and packet size, and is based on and helps determine the QoS. These differential service classes help network service providers manage the flow of mission-critical data through WANs using Cisco IOS[®] Software.

Quality of Service Overview

Twenty-five years ago, QoS was an abstract term more related to practical considerations than technical factors. Mainframes and client-server systems handled batch processes and basic data transactions with little or no real-time demands. All this changed with the advent of the Internet and advanced applications, including IP telephony, IP videoconferencing, multicasting, and enterprise software systems that are ineffective without adequate performance levels.

Performance has improved as QoS capabilities have evolved along with other advancements such as load balancing, mirroring, and caching. With this evolution, however, came increased complexity for continually configuring, monitoring, and adjusting bandwidth resources.

Cisco IOS Software—the operating system software for Cisco networking products—provides tools that enable a higher standard for QoS. Through a combination of monitoring and control panels, network administrators can adjust performance for bandwidth and hardware as needed. QoS is the measure of performance that reflects the transmission quality and service availability of a transmission system. Service availability is a crucial foundational element of QoS. Before any QoS can be implemented successfully, the network infrastructure must be designed to be highly available. (The target for high availability is 99.999 percent uptime, with only five minutes of downtime permitted per year.) The transmission quality of the network is determined by the following factors:

- *Availability*—The fraction of time that network connectivity is available between an ingress point and a specified egress point is defined as network availability. Service availability is defined as the fraction of time that service is available between a specified ingress point and a specified egress point within the bounds of a defined service-level agreement (SLA).
- *Loss*—A comparison of the number of packets successfully transmitted and received to the total number of packets that were transmitted. Loss is expressed as the percentage of packets that were dropped. Loss is typically a function of availability. If the network is highly available, then loss (during periods of noncongestion) would be zero. During periods of congestion, however, QoS mechanisms determine which packets are suitable to drop.
- *Delay*—The amount of time it takes a packet to reach the receiving endpoint after being transmitted from the sending endpoint. In the case of voice, this delay is defined as the amount of time it takes for sound to leave the speaker's mouth and be heard in the listener's ear.
- *Delay variation (jitter)*—The difference in the end-to-end delay between packets. For example, if one packet requires 100 milliseconds (ms) to traverse the network from the source endpoint to the destination endpoint, and the following packet requires 125 ms to make the same trip, then the delay variation is calculated as 25 ms.
- *Throughput*—The available user bandwidth between an ingress point of presence (POP) and an egress POP.

Each end station in a VoIP or video over IP conversation has a *jitter buffer*. Jitter buffers are used to smooth out changes in arrival times of data packets containing voice. A jitter buffer can be dynamic and adaptive, and some Cisco codec can adjust for up to a 30-ms average change in arrival times of packets. If there are instantaneous changes in arrival times of packets that are beyond the jitter buffer's ability to compensate, there will be jitter buffer overruns and underruns, both of which result in an audible degradation of call quality.

In particular, QoS features provide better and more predictable network service by the following methods:



- Supporting dedicated bandwidth
- Improving loss characteristics
- Avoiding and managing network congestion
- Shaping network traffic
- Setting traffic priorities across the network

The IETF defines the following two architectures for QoS:

- Integrated Services (IntServ)
- Differentiated Services (DiffServ)

CoS is the process used to mark and separate packets into multiple priority levels. This distribution function enables simultaneous performance and service scalability for large-scale support of differentiated services.

What is the Cisco QoS Toolset?

Cisco Systems® provides a complete toolset of QoS features and solutions for addressing the diverse needs of voice, video, and data applications. Cisco QoS technology within Cisco IOS Software lets complex networks control and predictably service a variety of networked applications and traffic types. Bandwidth, delay, jitter, and packet loss can be effectively controlled. By helping ensure the desired results, the QoS features lead to efficient, predictable services for business-critical applications.

AutoQoS, a new capability in Cisco IOS Software, dramatically simplifies QoS deployment by automating Cisco IOS QoS features for voice traffic in a consistent fashion and using the advanced capability and intelligence of Cisco IOS Software. Cisco AutoQoS provides the user a simple, intelligent command-line interface (CLI) for enabling campus LAN and WAN QoS for Cisco VoIP switches and routers. The network administrator does not need extensive knowledge of the underlying network technology (Point-to-Point Protocol [PPP], Frame Relay, ATM, ATM-to-Frame Relay internetworking), required QoS service policies, or link-efficiency mechanisms needed to ensure voice quality and reduce latency, jitter, and packet drops.

AutoQoS capability automates network QoS, or packet prioritization and delivery. This capability benefits companies of all sizes by helping them to more easily deploy and manage QoS for IP and reduce operating costs.

AutoQoS simplifies implementation of IP QoS in LANs and WANs. With AutoQoS, Cisco automates the IP infrastructure to implement VoIP and IP communications from the wiring closet, across large enterprise IP backbone networks, and for service-provider-managed services. Small and medium-sized businesses benefit from the ability to more easily implement IP QoS in their networks and reduce their operating costs. This level of automation also allows enterprise companies, which might not have the required staffing and resources, to implement QoS, and helps service providers that must enable QoS on hundreds to thousands of customer-premises devices as part of their managed-service offerings, to do so with greater ease. With this feature, QoS deployments can be faster and up to three times less expensive than before.

Classification and Marking Tools

The first requirement of a QoS policy is to identify the type of traffic that requires different treatment (either preferentially or deferentially). Classification tools mark a frame or packet with a specific value. This marking (or re-marking) establishes a trust boundary upon which scheduling tools, such as Class-Based Weighted Fair Queuing (CBWFQ) and Modified Deficit Round Robin (MDRR) queuing, later depend.

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

- Layer 2 parameters (802.1Q CoS bits, Multiprotocol Label Switching experimental values [MPLS EXP])



- Layer 3 parameters (IP Precedence, differentiated services code point [DSCP], source or destination IP address)
- Source port, destination port, or stateful inspection

QoS policies can be applied to traffic only after it is positively identified. Best-practice design recommendations are to identify and mark traffic (with DSCP values) as close to its source as possible. The network edge where markings are accepted (or rejected) is referred to as the “trust boundary.” If markings and trusts are set correctly, then intermediate hops do not have to perform detailed traffic identification, but instead can administer QoS policies (such as scheduling) based on these previously set DSCP markings. This approach simplifies and modularizes QoS policy administration and reduces the CPU overhead of the router required to enforce QoS policies.

There are several mechanisms that can be used for marking traffic, including:

- *802.1Q/p CoS*—Only three bits are available for 802.1p marking, so only eight classes of service (0–7) can be marked on Layer 2 Ethernet frames.
- *IP type of service (ToS) byte*—The second byte in an IPv4 packet is the ToS byte. The first three bits of the ToS byte alone are referred to as the IP Precedence bits. The IP Precedence bits, like 802.1p CoS bits, allow for only eight values of marking (0–7). Many enterprises find IP Precedence marking to be overly restrictive and limiting based on the number of classes available, favoring instead the 6-bit, 64-value DSCP marking model.
- *DSCPs and Per-Hop Behaviors (PHBs)*—DSCP values can be expressed in numeric form or by special keyword names, called Per-Hop Behaviors (PHBs). There are three defined classes of DSCP markings: best effort (BE or DSCP 0), assured forwarding PHBs (AFxy), and expedited forwarding (EF). In addition to these three defined PHBs, there are class selector code points that have been defined to be backward-compatible with IP Precedence (CS1-CS7, which are identical to IP Precedence values 1–7).

There are four assured-forwarding classes, denoted by the letters “AF” followed by two numbers.

- *MPLS EXP*—MPLS EXP bits are the three bits within the MPLS label that are used to hold a QoS indicator, which by default is copied down from the IP Precedence field in the underlying IP packet during label imposition. This field allows up to eight different QoS markings, versus 64 for DSCP. These EXP bits are used to determine the PHB for the MPLS nodes and can also be used as transparency mechanisms when used with MPLS DiffServ Tunneling Modes, such as Pipe and Uniform Modes. More information on MPLS DiffServ Tunneling Modes and how they can achieve customer QoS transparency are discussed in the QoS Transparency section of this paper.

Scheduling Tools

Scheduling tools determine how a frame or packet exits a device. Whenever packets enter a device faster than they can exit it (as with speed mismatches), a bottleneck can occur. Devices have buffers that allow for scheduling higher-priority packets to exit sooner than lower-priority ones, called queuing. Queuing algorithms are activated only when a device is experiencing congestion, and in most cases are deactivated when the congestion clears.

Queuing buffers are finite in capacity and act very much like a liquid pouring into a container through a funnel. If water is continually entering the funnel much faster than it exits, eventually the funnel is overflowing from the top. When queuing buffers begin to overflow from the top, packets are dropped—either as they arrive (tail-drop), or selectively before all buffers are filled. Selective dropping of packets during packet queuing 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 flow to manageable rates.



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

Scheduling tools include: CBWFQ, MDRR Queuing, Low-Latency Queuing (LLQ), and Weighted Random Early Detection (WRED).

Link-Specific Tools

Link-specific tools include policing and shaping tools, link fragmentation and interleaving (LFI) tools, compression tools, and transmit ring tuning.

Policers and shapers usually identify traffic violations in an identical manner; their main difference is the manner in which they respond to violations. A policer typically drops traffic, whereas a shaper typically delays excess traffic using a buffer to hold packets and shape the flow when the data rate of the source is higher than expected.

LFI for multilink PPP and Frame Relay fragmentation are the two main tools used to mitigate serialization delay on slow links. (With slow-speed WAN circuits, large data packets take an excessively long time to be placed onto the wire. This delay is referred to as serialization delay, and can easily cause a VoIP packet to exceed delay and/or jitter threshold.)

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 very large and can account for nearly two-thirds of the entire packet. To avoid the unnecessary consumption of available bandwidth, CRTP can be used on a link-by-link basis. CRTP compresses IP, User Datagram Protocol (UDP), and Real Time Protocol (RTP) headers from 40 bytes to 2–5 bytes.

The transmit ring is a final FIFO queue that holds frames to be immediately placed on to the physical interface. Its purpose is to ensure that a frame will always be available when the interface is ready to transmit traffic, so that link usage is increased to 100 percent of capacity. The size of the transmit ring depends on the hardware, software, Layer 2 media, and queuing algorithm configured on the interface. It is a general best practice to set the transmit ring to a value of three on slow-link interfaces.

Enterprise QoS Requirements and the QoS Baseline

When designing a network, service providers need to recognize what the enterprise QoS requirements are in order to meet these customer needs.

The Cisco QoS feature set presents myriad deployment options and combinations—and nearly every QoS-savvy engineer has a slightly different opinion on how best to enable them. To present a consistent QoS story, Cisco has adopted a new initiative called the “QoS Baseline,” designed to unify QoS implementation on Cisco platforms.

The QoS Baseline specifies the default platform marking and behavior for (up to) 11 traffic classes within the enterprise. These are described in more detail later. Note that the QoS Baseline considers the QoS needs of today as well as the foreseeable future by providing for 11 traffic classes. Even if an enterprise needs to provision for only a handful of these 11 classes today, following QoS Baseline recommendations will enable them to leave options open for smoothly provisioning additional traffic classes in the future.

The 11 traffic classes are: routing, voice, interactive voice, streaming video, mission-critical data, call signaling, transaction data, network management, bulk data, scavenger, and best effort.



AutoQoS in its second version will automatically configure QoS for voice, video, and data in an enterprise environment. Cisco AutoQoS Enterprise will detect and provision for up to 10 classes of traffic, based on the QoS Baseline. (The only class not automatically provisioned will be locally defined mission-critical, because this class requires a business-level awareness that is beyond the tool's capabilities; this business-level factor will be discussed further in the "Locally-Defined Mission-Critical Class" section.) AutoQoS Enterprise is targeted to abstract and simplify the complexity of managing a QoS Baseline-compliant design. Although AutoQoS is not relevant to the service provider's core network, it is important for a service provider to understand the implications that AutoQoS has upon enterprise customer networks. As AutoQoS becomes more widespread within an enterprise network, service providers will see the need for QoS deployments to become accelerated.

QoS Requirements for Voice

When considering the QoS needs of enterprise VoIP traffic, remember that voice quality is directly affected by all three QoS quality factors: loss, delay, and delay variation.

Loss causes voice clipping and skips. The industry-standard codec algorithms used in Cisco digital signal processors (DSPs) can correct for up to 30 ms of lost voice with the use of concealment algorithms; therefore, the loss of two or more consecutive 20-ms voice samples will result in noticeable degradation of voice quality. Assuming a random distribution of drops within a single voice flow, a drop rate of just 1 percent in a voice stream will result in a loss that cannot be concealed once every 3 minutes, on average; a 0.25 percent drop rate results in a loss that cannot be concealed once every 53 minutes, on average.

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

Regarding *delay variation*, there are adaptive jitter buffers within Cisco IP Telephony devices that can usually only compensate for 20–50 ms of jitter. These jitter buffers are dynamically adaptive, so there is no defined and absolute limit for jitter that will hold true for all circumstances. However, testing has shown that when jitter consistently exceeds 30 ms, then voice quality degrades significantly.

In centralized call-processing designs, IP phones use a TCP control connection to communicate with the Cisco CallManager. If not enough bandwidth is provisioned for these lightweight control connections, the user might be adversely affected. An example is the delay-to-dial-tone time periods. When an IP phone goes off-hook, it "asks" the Cisco CallManager what to do. The Cisco CallManager instructs the IP phone to play a dial-tone. If control traffic is dropped or delayed within the network, the user will not get the dial-tone, which he is expecting immediately. This same logic applies to all signaling traffic for gateways and phones.

For Cisco IP phones, the control traffic required is approximately 150 bits per second (bps) per phone (not including Layer 2 overhead).

QoS Requirements for Data

When addressing the QoS needs of data application traffic, no more than four main traffic classes should be used. It is important to profile applications to get a basic understanding of their network requirements and traffic patterns, but the network should not be overly engineered. Typical CoS classes might include:



- *Locally Defined Mission-Critical*—Transactional and interactive applications with a high business priority
- *Transactional/Interactive*—Client-server applications, messaging applications
- *Bulk*—Large file-transfers, e-mail, network backups, database syncs and replication, video content distribution
- *Best Effort*—Default class for all unassigned traffic; provision at least 25 percent of bandwidth as Best Effort

An optional (deferential) class is *Scavenger*, used for peer-to-peer media-sharing applications, gaming traffic, and entertainment traffic. Additional optional classes include *Routing* and *Network Management*.

Best Effort Class

The *Best Effort* class is the default class for all data traffic. An application will be removed from the default class only if it has been selected for preferential or deferential treatment. Because many enterprises have hundreds, even thousands, of data applications running over their networks (the majority of which will remain assigned to this default class), adequate bandwidth needs to be provisioned for the default class. It is recommended that at least 25 percent of a WAN link's bandwidth be reserved for the default Best Effort class.

Bulk Data Class

The *Bulk Data* class is intended for applications that are relatively noninteractive and drop-insensitive, which 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 where users are *not* typically unable to proceed because they are waiting for the completion of the operation.

The advantage of provisioning bandwidth to Bulk Data applications (rather than applying policing policies to them) is that these applications can dynamically take advantage of unused bandwidth and thus speed up their operations during nonpeak periods, which reduces the likelihood of their leaking into busy periods and absorbing inordinate amounts of bandwidth for their time-insensitive operations.

Transactional/Interactive Data Class

The *Transactional/Interactive* class is a combination of two similar types of applications: transactional client-server applications and interactive-messaging applications. The response-time requirement separates transactional client-server applications from generic client-server applications. With transactional client-server applications (such as SAP, PeopleSoft, and DLSw+), the user is waiting for the operation to complete before proceeding. E-mail is not considered a transactional client-server application, because most e-mail operations happen in the background and users usually do not notice even several hundred ms delays in mailspool operations.

Note: By default, DLSw+ (a transactional application) marks its traffic to IP Precedence 5, which interferes with VoIP; therefore, it is recommended to re-mark DLSw+. This is detailed in the “QoS in an Cisco AVVID-Enabled Wide-Area Network” chapter of the *Cisco AVVID Network Infrastructure Enterprise Quality of Service Design*.

Locally Defined Mission-Critical Data Class

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, namely for each enterprise to have a premium CoS for a select subset of its transactional and interactive applications that have the highest business priority for *that business*. For example, an enterprise may have properly provisioned Oracle, SAP, BEA, and DLSw+ within their



Transactional/Interactive class. However, the majority of its revenue may come from SAP, and therefore it may want to give this transactional application an even higher level of preference by assigning it to a dedicated class (such as a Locally-Defined Mission-Critical class).

Scavenger Class

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, etc.), gaming applications (Doom, Quake, Unreal Tournament, etc.), and any entertainment video applications. This is a typical class defined for the enterprise, but is typically re-marked with the Best Effort class at the service provider edge.

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

Routing and Network Management Classes

Some enterprises choose to explicitly provision a minimal bandwidth queue for routing and other network-control applications (such as IP Security [IPSec] traffic). Similarly, a separate minimal bandwidth queue can be provisioned for network-management traffic, which could include Simple Network Management Protocol (SNMP), Network Time Protocol (NTP), Syslog, and Network File System (NFS) traffic.

Note: It is important to note that Interior Gateway Protocol (IGP) traffic (such as Routing Information Protocol [RIP] and Enhanced Interior Gateway Routing Protocol [EIGRP]) typically does not require explicit traffic provisioning, because it benefits from the Cisco internal mechanism of PAK_PRIORITY. Of note, within Open Shortest Path First (OSPF) Protocol only the hellos are marked with the PAK_PRIORITY, and Border Gateway Protocol (BGP) traffic (while also marked IPP6/CS6) does not receive such preferential treatment and may need to be explicitly protected to maintain peering sessions. For more information on PAK_PRIORITY please visit:

<http://www.cisco.com/warp/public/105/rtgupdates.html>

Guidelines for Collapsing Enterprise Classes

Although Cisco is adopting its new QoS Baseline initiative and designing tools like Cisco AutoQoS Enterprise to facilitate and simplify the deployment of complex QoS traffic models within the enterprise, to date very few enterprises have deployed more than a handful of traffic classes. Therefore, most service providers offer only a limited number of classes within their MPLS VPN. At times, this may require enterprises to collapse the number of classes they have provisioned to integrate into their service provider’s QoS models. The following factors must be considered when deciding how best to collapse and integrate enterprise classes into various service provider QoS models.

Voice and Video

Service providers typically offer only one “real-time” class or “priority” class of service. If an enterprise wants to deploy both voice and interactive video (and it is recommended that both be provisioned with strict-priority treatment) over its MPLS VPN, then it has difficulty deciding which one should be assigned to the real-time class, and whether there are any implications about assigning both to the real-time class.



Voice and video should never both be assigned Low-Latency Queuing (LLQ) on link speeds where serialization is a factor (less than or equal to 768 kbps). Packets offered to the LLQ are not typically fragmented, and thus large IP videoconferencing packets may cause excessive delays for VoIP packets on slow links.

An alternative may be to assign IP videoconferencing to a nonpriority class, which entails not only the obvious caveat of lower service levels, but also possible traffic-mixing considerations, such as call signaling and mixing TCP with UDP.

VoIP requires provisioning not only of RTP bearer traffic but also call-control or -signaling traffic, which is lightweight and only requires a moderate amount of guaranteed bandwidth. Because the service levels applied to call-signaling traffic directly affect delay-to-dial-tone, it is important from the end-user's expectations that call signaling be protected. Service providers may not always offer a suitable class just for call-signaling traffic by itself, so the question arises as to which other traffic classes call signaling should be mixed with.

On links where serialization is not an issue (greater than 768 kbps), call signaling can be provisioned into the real-time class, along with voice. This is not recommended on slower links—rather, assign call signaling into one of the preferential data classes for which the service provider provides a bandwidth guarantee. It is important to realize that a guarantee applied to a service provider class as a whole does not in itself guarantee adequate bandwidth for an individual enterprise application.

It is a general best practice not to mix TCP-based traffic with UDP-based traffic (especially streaming video) within a single service provider class, due to the behaviors of these protocols during periods of congestion. Specifically, TCP transmitters will throttle-back flows when drops have been detected. Although some UDP applications have application-level windowing, flow control, and retransmission capabilities, most UDP transmitters are completely oblivious to drops and thus never lower transmission rates due to dropping. When TCP flows are combined with UDP flows within a single service provider class and the class experiences congestion, then TCP flows will continually lower their rates, potentially giving up their bandwidth to drop-oblivious UDP flows. This effect is called TCP-starvation/UDP-dominance.

Even if Weighted Random Early Detection (WRED) is enabled on the service provider class, the same behavior occurs, because WRED usually affects only TCP-based flows.



Marking Considerations

Service providers use Layer 3 marking attributes (IP Precedence or differentiated services code point [DSCP]) to determine which CoS to assign to the packets they receive. Therefore, enterprises must mark or re-mark their traffic consistent with their service provider's admission criteria to gain the appropriate level of service. Additionally, service providers may re-mark out-of-contract traffic within their network cloud, which may affect enterprises that require consistent end-to-end markings. The following points should be considered when determining an enterprise-to-service provider marking and re-marking strategy.

Enterprise-to-Service Provider Re-marking

A general enterprise-marking rule is to mark or trust traffic as close to the source as administratively and technically possible. However, certain traffic types may need to be re-marked before handoff to the service provider to gain admission to the correct class. If such re-marking is required, it is recommended that the re-marking be performed at the content engine's egress edge, and not within the campus network, because service provider service offerings will likely evolve or expand over time, and adjusting to these changes will be easier to manage if such re-marking is performed only at the content engine's egress edge.

When multiple types of traffic must be marked to the same code-point value to gain admission to the appropriate queue, such as on high-speed links, it may be preferable to send voice, interactive video, and call signaling to the service provider's real-time class. If this service provider class only admits DSCP EF and CS5, then two of these three applications would be required to share a common code point.

Service Provider-to-Enterprise Re-marking

Service providers may re-mark traffic at Layer 3 to indicate whether certain flows are out of contract. This is consistent with DiffServ standards, such as RFC 2597. However, certain enterprises require consistent end-to-end marking, typically for management or accounting purposes. In such cases the enterprise may choose to apply re-marking policies as traffic is received back from the service provider's MPLS VPN (on the *ingress* direction of the enterprise's content engine).

Class-based marking can again be used because it supports not only access lists for classification, but also Network-Based Application Recognition (NBAR).

QoS Transparency with MPLS DiffServ Tunneling Modes

It is often preferable for the service provider to maintain its own QoS service policies and customer SLAs without overriding the enterprise customers' own DSCP or IP Precedence values. MPLS can be used to tunnel a packet's QoS markings and create QoS transparency for the customer. For example, it is possible to mark the MPLS experimental values (MPLS EXP) field differently (even independently) of the Per-Hop Behavior (PHB) marked in the IP Precedence or DSCP fields. A service provider may choose from an existing array of classification criteria, including or excluding the IP PHB marking, to classify those packets into a different PHB, which is then marked only in the MPLS EXP field during label imposition. This is useful, for example, to a service provider that requires SLA enforcement of its customer's packets by promoting or demoting a packet's PHB without regard to the customer's QoS marking scheme and without overwriting the customer's IP PHB marking. This can be thought of as adding a layer of PHB to a packet or encapsulating the packet's PHB with a different QoS Tunnel PHB layer. There are three distinct MPLS DiffServ tunneling modes (which are described in RFC 3270): uniform mode, short pipe mode, and pipe mode.

Uniform mode is used when the customer and service provider share the same DiffServ domain. The outermost header is always used as the single meaningful information source as it relates to the QoS PHB. On MPLS label imposition, the IP Precedence classification is copied into the outermost label's EXP field. This is the default behavior.



Short pipe mode is used when the customer and service provider are in different DiffServ domains. This is useful when the service provider wants to enforce its own DiffServ policy and the customer requests that the customer DiffServ information be preserved, thus providing a DiffServ transparency through the service provider network. The outmost label is used as the single meaningful information source as it relates to the service provider's QoS PHB.

Pipe mode is very similar to short pipe mode because the customer and service provider are in different DiffServ domains. The difference between the two is that with pipe mode, the service provider derives the outbound classification for WRED and WFQ based on the service provider's own DiffServ policy (rather than according to the enterprise customer's markings). This affects how the packet is scheduled on the egress provider edge prior to the label being popped.

Enterprise-to-Service Provider Mapping Models

Service providers may offer multiple QoS models for their MPLS VPN services. An enterprise would not likely have more than 3–5 models on T1 or lower links, so a 1:1 mapping may be adequate. On higher-speed links, mapping into fewer service provider classes may be a necessity.

- *Three-Class Service Provider Model*—In this model, the service provider offers three classes of service: real time (strict priority), critical data (guaranteed bandwidth), and best effort. Under such a model, there is no recommended provision for protecting streaming video (following the “Don’t mix TCP with UDP” guideline), nor is there a service provider class suitable for bulk data, which consists of large, non-“bursty” TCP sessions which could drown out smaller data transactions.
- *Four-Class Service Provider Model*—Building on the previous model, a fourth class is added which may be used for either bulk data or streaming video.
- *Five-Class Service Provider Model*—Building again on the previous model, a fifth class is added which may also be used for either bulk data or streaming video (whichever was not used under the four-class model).

Service Provider QoS Requirements

Service Provider SLA Requirements

End-to-end QoS is like a chain, which is only as strong as the weakest link; therefore, it is essential for enterprises to use service providers that can provide the SLAs required for Cisco AVVID (Architecture for Voice, Video and Integrated Data) applications. For example, the end-to-end SLA requirements of voice and interactive video are: no more than 1 percent loss, no more than 150 ms of one-way latency from mouth to ear (per ITU G.114 standard), and no more than 30 ms jitter.

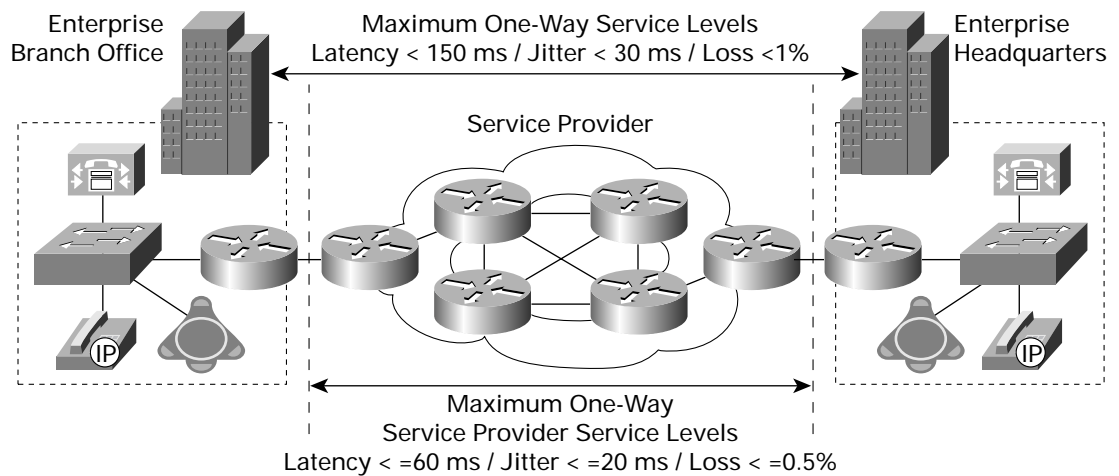


Thus, the service provider's component (a subset of the trip) must be considerably tighter. Figure 1 shows the following SLAs for IP multiservice as defined for Cisco Powered Network.

- No more than 0.5 percent loss
- No more than 60 ms of one-way latency from edge to edge
- No more than 20 ms of jitter

Figure 1

Cisco Powered Network Multiservice Service Provider SLAs



To achieve such end-to-end SLAs, enterprises and service providers must cooperate and be consistent in classifying, provisioning, and integrating their respective QoS solutions.

Service Provider Backbone Network Design Considerations

Several options exist to meet strict SLA considerations for loss, delay, and jitter in the service provider backbone network:

- *Aggregate bandwidth overprovisioning*—This is a common trend in the service provider backbone network due to its simplicity and ease to design, deploy, and operate. DiffServ domain characteristics are assumed at the edge for aggregation of traffic. Studies have shown that designing the service provider backbone network for low delay, jitter, or loss can simply be a matter of overprovisioning the network by approximately two times the maximum of the aggregate traffic load. Caveats to overprovisioning include: capacity planning failures, network failure situations, and unexpected traffic demands or patterns. Also, in this instance there is no differentiation between Priority-Queuing class traffic and Best Effort, so in the event of failure or congestion the Priority-Queuing traffic can be degraded. This method may also not provide the most cost-effective solution.
- *DiffServ in the backbone network*—Deploying a modest DiffServ policy in the backbone network allows the service provider to support multiple classes of traffic with different overprovisioning and underprovisioning ratios on a per-class basis. DiffServ in the backbone network allows for two cases of traffic conditions: less bandwidth is required to achieve the same SLA when compared to non-DiffServ case, *or* it can be assumed that more aggregate traffic is supported for the same provisioned network bandwidth as the non-DiffServ backbone network. The caveats to this solution are adding complexity to the network design and operations. In a DiffServ backbone network, it may not be necessary to assume the



same number of classes that exist at the edge in the provider edge-to-customer edge link, so long as the classes that are defined assume an Express-Forwarding class for real-time traffic (voice and video) associated to a priority queue and critical data is protected.

DiffServ Backbone-Network Example—Three-Class Backbone-Network Model

It is not necessary to ensure that the backbone network supports the same number of DiffServ classes as the edge, assuming that proper design principles are in place to support the given SLAs. One example of this is to provision three DiffServ classes in the backbone network, while five classes are provisioned at the provider edges, as shown in Table 1.

Table 1 QoS Standard Classification and Marking Recommendations

DiffServ Provider-Edge Classes	DiffServ Backbone-Network Classes
Real Time	Core Real Time
(Streaming) Video	
Critical Data	Core Critical Data
Bulk	
Best Effort	Core Best Effort

Backbone-Network Classes Definitions

Backbone-network classes are defined as follows:

- *Core Real Time*—This class targets applications such as VoIP and interactive video, which require low loss (less than 0.25 percent), low delay, and low jitter (typically 5 ms within the backbone), and have a defined availability. This class may also support per-flow sequence preservation. This class should always be engineered for the worst-case delay to support the real-time traffic. Excess traffic in this class is typically dropped. This class should be associated to expedited forwarding with a priority queue to ensure that the delay and jitter contracts are met. Between 25–33 percent of link capacity should be allocated to the priority queue. WRED should not be configured on this queue.
- *Core Critical Data*—This class represents business-critical interactive applications such as Systems Network Architecture (SNA), SAP R/3, Telnet, and possibly intranet Web applications to selected URLs. It is defined in terms of delay (round-trip time [RTT] should be less than 250 ms—the threshold for human delay perception) and loss (less than 1 percent loss rate is typical, with targets as low as 0.1 percent also available), with an availability. Throughput is derived from loss and RTT. Jitter is not important for this service class and is not defined. Excess in this class is typically re-marked with an out-of-contract identifier (re-marking of EXP to a lower value) and transmitted. This class may also support per-flow sequence preservation. This class should be associated with an assured-forwarding class-based queue and assigned up to 90 percent of the remaining bandwidth (once the priority-queuing and expedited-forwarding traffic has been serviced). WRED should be configured here to optimize TCP throughput and to accommodate a drop policy for out-of-profile traffic.
- *Core Best Effort*—This class represents all other customer traffic that has not been classified as *Real Time* or *Critical Data*. It is defined as a loss rate with availability; throughput is derived from loss. Delay and jitter are not important for this service and are not defined; therefore, only 10 percent of remaining link capacity (after the priority queue has been served) should be allocated to this queue.

On the service provider's edge router, a mapping capability is thus essential to map several edge classes into a single aggregate backbone-network class. In the previous example, several provider-edge classes (streaming video, critical data, and bulk data) are mapped into a single backbone-network class (Core Critical Data). This mapping can be realized one of two ways:

- *A backbone-network class matches several DSCPs*—Applying this to the previous example, if DSCP AF31 represents critical data at the edge, DSCP AF21 represents (streaming) video at the edge, and DSCP AF11 represents bulk data at the edge, then the backbone-network aggregate class (Core Critical Data) matches on DSCPs AF31, AF21, and AF11.
- *When MPLS is used in the backbone network, the edge service provider router can set the MPLS EXP field (3 bits) as a function of the received DSCP*—Applying this to the same example, if MPLS EXP=3 is used for the backbone-network aggregate class (Core Critical Data), the service provider's edge routers will impose MPLS labels with EXP=3 for packets received with DSCP AF31 (edge critical data), AF21 (edge streaming video), or DSCP AF11 (edge bulk data).

Summary

Service providers that use important QoS components in network planning can provide a greater value to the enterprise customer. Service providers are then able to reduce the expenditures of providing more bandwidth, and at the same time provide the enterprise customer with the tight SLAs essential to services such as VoIP and interactive video, meeting and exceeding enterprise demands for QoS and CoS for voice, data, video, etc. Service provider offerings need to have strong QoS and CoS to win and keep enterprise business.

By deploying specific QoS mechanisms to manage delay, delay variation (jitter), bandwidth, and packet loss on a network, service providers can achieve the level of end-to-end QoS that their enterprise customers require.

A communications network forms the backbone of any successful organization. These networks serve as a transport for a multitude of applications, including delay-sensitive voice, and bandwidth-intensive video. These business applications stretch network capabilities and resources, but also complement, add value, and enhance every business process. Networks must therefore provide secure, predictable, measurable, and sometimes guaranteed services to these applications. Achieving the required QoS—by managing the delay, jitter, bandwidth, and packet loss parameters on a network, while maintaining simplicity, scalability, and manageability, is the secret to running an infrastructure that truly serves the business and the enterprise.



Corporate Headquarters

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

European Headquarters

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

Americas Headquarters

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

Asia Pacific Headquarters

Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

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

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