How Cisco IT Builds End-to-End QoS into Its Network

QoS Version 2 yields consistent management and performance standards across LAN and WAN.

Cisco IT Case Study / Network Management / End-to-End QoS in the Network: This case study describes Cisco IT's internal deployment of quality of service within the Cisco network, a leading-edge enterprise environment that is one of the largest and most complex in the world. Customers can draw on Cisco IT's real-world experience in this area to help support similar enterprise needs.

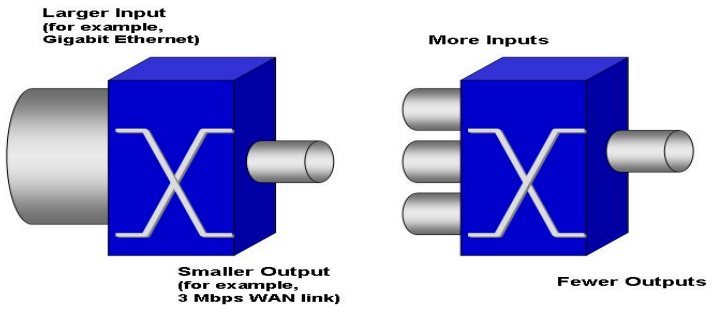
> "Quality of service is absolutely indispensable to the support of real-time voice and video applications and [is] critical to network capacity planning. Without it, Cisco could potentially experience severe degradation of network performance and be unprepared to deploy the enterprise applications that are emerging."
>
> **Shawn Yapa, Global IT Project Manager, Cisco**

BACKGROUND

In 1995, the Cisco IT team began a link-by-link implementation of quality of service (QoS), a vital element of network operations. By the late 1990s, the emergence of voice and video applications added urgency to the company's efforts to categorize and control its network traffic. These latency-sensitive, bandwidth-intensive applications add value to Cisco business processes, but they also tend to strain network capabilities and resources. Despite the belief of some IT professionals that networks built with high-capacity switches and high-speed LAN and WAN links do not require QoS, the technology has evolved into a necessity.

All network topologies have points where congestion builds and data can be lost: WAN links where a large trunk or several trunks funnel into one (Figure 1). It is useful to view QoS as a basic foundation technology in deploying or upgrading networks that carry real-time voice, video, and business applications.

Figure 1. QoS: A Requirement When Network Congestion Occurs



CHALLENGE

By 2003, Cisco IT—propelled by both business and technology drivers—was ready to develop a second version of QoS, referred to internally as QoS V2. "We needed a comprehensive, end-to-end solution that addressed the network consistently across user devices from the LAN edge to the enterprise WAN edge," says Liem Nguyen, Cisco IT network engineer and team lead for global QoS design. "We wanted to standardize and consolidate QoS on these

networks. Previous deployments of QoS were inconsistent. Perhaps the most immediate driver for QoS V2 was the impending deployment of our Unity voicemail system—we knew we could not effectively support real-time applications, particularly voice, when we were operating with many unique, isolated environments. ”

Cisco was experiencing other challenges as well:

- Traffic from Cisco labs required the prioritization QoS provides, yet this traffic could be problematic, given the labs’ development, test, and integration environment.
- Employees who worked from home offices frequently required IP voice capabilities (IP telephony) over Cisco’s virtual private network (VPN) and expressed concern over service quality.
- Communications and network service providers handled different types of network traffic (classes of service) differently and were not always aligned with the way Cisco classified network traffic (Table 1). This frequently resulted in packet drop, delay, and delay variation (jitter).

Table 1. Cisco IT QoS Classes of Service

QoS Class	Description	IP Precedence	DSCP
6	Network control/routing traffic	6 and 7	48–63
5	Voice bearer traffic	5	46
4	Videoconferencing bearer traffic	4	32–39
3	Signaling	3	24–31
2	High-priority data	2	16–23
1	Batch/scavenger traffic	1	8–15
0	Default/best-effort	0	0

“We were tackling several big issues simultaneously,” says Lee Packer, the Cisco project manager who supervised the QoS deployments for Cisco’s Americas and Asia Pacific regions. “We needed to respond to user demand for applications like desktop videoconferencing. We needed to support a very high volume of data backup and replications and other traffic across our LAN during periods of heavy congestion. We also needed a foundation strategy that allowed for the integration of voice, video, and Call Admission Control that could intelligently handle oversubscribed call requests.”

“It all came down to trust,” adds Nguyen. “To offer an end-to-end QoS solution, we needed a QoS architecture that would give us the freedom to manipulate and manage traffic only at the network LAN edge, while implicitly trusting all traffic within the core network. In essence, we wanted to create a QoS trust boundary at the Layer-2 LAN switches that would allow traffic to be correctly prioritized at the WAN edge without having to classify and mark traffic at every network devices in between. Simple to say, but not necessarily simple to do.”

SOLUTION

Keeping in mind its primary goal of building QoS based on trusted network edges, the Cisco IT design team established other requirements for QoS V2, including:

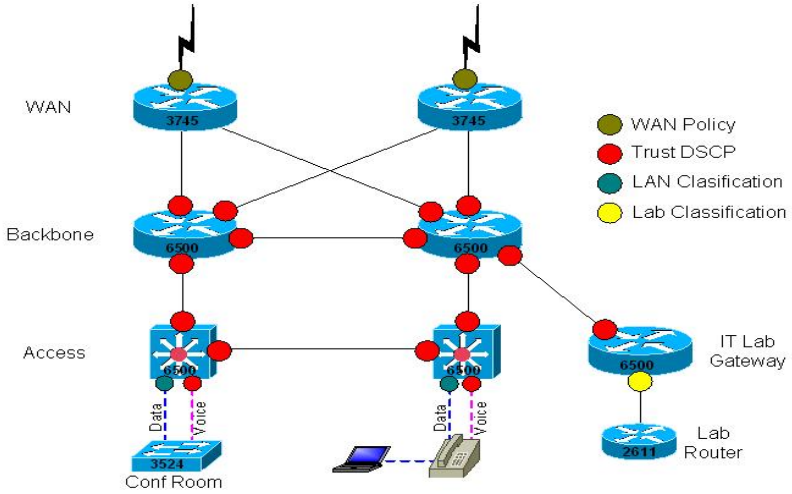
- Consistent performance and management standards across the LAN and WAN networks
- The lowest-possible latency for voice applications
- The ability to limit bandwidth consumption by large data transfers during times of traffic congestion

- Reserved bandwidth for real-time voice and video, critical business applications, and signaling
- Class-Based Weighted-Fair Queuing (CBWFQ) policy roadmaps that scale to meet the needs of a global network
- Low Latency Queuing (LLQ) to add a strict priority queue to the CBWFQ scheduler and to handle voice and interactive-video applications without performance degradation
- The ability to standardize and minimize device configurations to ensure consistent results across all Cisco locations

QoS V2 Design Strategy Overview

The design strategy driving a consistent end-to-end network implementation of QoS V2 was relatively straightforward. The Cisco QoS design team decided that all traffic should follow clearly defined processes as it proceeded from source to destination devices. The team established the trust/classification boundary for traffic at the LAN edge closest to connected devices they considered external to the Cisco network infrastructure—desktops, IP telephony devices, servers, or Cisco lab gateways. Using this topology, all IP traffic arriving at the LAN edge switch was classified as trusted or untrusted. Trusted traffic entered the network with the Differentiated Services Code Point (DSCP) values unchanged. The DSCP value is rewritten to zero for all untrusted traffic, except for a known list of applications that matches specific User Datagram Protocol (UDP) and Transmission Control Protocol (TCP) ports, which are remarked to the appropriate priority that matches the Cisco IT classes of service defined in Table 1. The team defined untrusted traffic as inbound packets from data VLANs, the Internet, remote-access users, and the majority of Cisco labs (Figure 2).

Figure 2. High-Level QoS Design



Traffic Classification and Marking

The Cisco IT class of service strategy makes it simple to group applications according to their required priorities, creating a model that is easy to maintain. Real-time voice bearer traffic is marked as Class 5 and guaranteed expedited delivery using LLQ to ensure that voice quality is not adversely affected during periods of high bandwidth use. Traffic classes 2, 3, 4, and 6 are assigned minimum bandwidth and queued by CBWFQ. Class 0 gets the remainder of the bandwidth, while Class 1 (batch/scavenger) traffic is slowed down during periods of heavy network use.

Special Design Considerations

During the course of developing QoS V2, the Cisco design team addressed several special design considerations,

including:

- **IP/VPN QoS strategy –** In the case of Layer-3 VPN technology such as Multiprotocol Label Switching (MPLS) VPN, Cisco classes of service may not map to those offered by IP service providers. As a result, agreed-upon service levels may not be met. To ensure that high-priority traffic is given the same level of service as it is when traversing the Cicso WAN, special care is necessary. To facilitate data handling and reduce voice quality and packet drop problems, the Cisco IT QoS team mandated that a service provider's network be able to accommodate a minimum of a real-time class (managed by LLQ), a high-priority class (managed by CBWFQ), a normal-priority class, and a best-effort class with a guaranteed delivery service-level agreement (SLA) of 99.9 percent.
- **Weighted Random Early Detection (WRED) –** This technique enables early detection of network congestion and selective discarding of lower-priority data packets based on their class of service values. For WRED to be successful, thresholds must be set correctly. If a threshold is too low, there may be unnecessary random data drops. This can cause voice and data applications to reduce their transmission rate, causing underutilization of bandwidth. However, if a threshold is set too high, traffic congestion can result. The Cisco IT QoS team believed that WRED could be helpful at Cisco WAN edge routers that connect directly to IP/VPN service providers—it would cause Cisco routers to discard traffic that might congest the core of the provider network.
- **Internet edge/VPN QoS strategy –** To extend QoS to Internet traffic that enters and leaves Cisco's network infrastructure, it is essential to classify the traffic without impeding the performance of Internet applications. If QoS is well-designed, outbound Internet traffic should be classified and marked at the LAN edge. Inbound Internet traffic should, at a minimum, be classified as 0 before entering the corporate firewall, with the exception of traffic destined for devices that have a higher-priority classification.

  As more enterprises adopt Internet VPNs for site-to-site connectivity and deploy remote-access solutions, it becomes increasingly important to ensure that supporting infrastructures can classify and prioritize traffic to and from the Internet. In addition, it is vital to be able to use software features such as QoS preclassification and support for LLQ in order to expedite high-priority data.
- **QoS for Cisco labs –** The Cisco IT QoS design team decided to classify traffic from Cisco labs as 1 (batch/scavenger) as close to the lab edge as possible. It also required labs that needed real-time video and voice capabilities to present a valid business justification for those needs. The plan was for Cisco IT and the labs to mutually define bandwidth requirements for each class of traffic requested. Any real-time traffic over agreed-upon limits would be policed to minimize impact on WAN resources.
- **Wireless networks –** Because the team considered wireless networks an extension of the wired network, it applied the same edge classification/marking strategy to them. Wireless data traffic is classified at the LAN edge via access control lists (ACLs). Voice data carried on virtual LANs is trusted on the access switch to which wireless network access points are connected. As wireless traffic continues to proliferate, the team plans to assess the QoS capabilities available to wireless access points, in order to preserve an end-to-end QoS solution.
- **QoS on backup WAN connections –** Cisco's internal WAN services are diverse; depending upon the geographical region, they may support a variety of technologies, including Asynchronous Transfer Mode (ATM) and network-based VPNs. Backup connections may offer lower bandwidth than primary circuits and, therefore, the potential for traffic congestion. A complicating factor is that the need to support real-time network traffic can vary significantly among regions. The design team established QoS criteria for backup circuits of all sizes in order to provide minimum bandwidth for traffic in higher-priority classifications.

RESULTS

Cisco IT QoS V2 has delivered several significant benefits. "One important result of our efforts is that there are significantly fewer voice-application-related calls," comments Shawn Yapa, global IT project manager. "Users are

pleased with the quality of voice and video, and helpdesk personnel are reporting few support problems specifically related to our Unity voicemail system." "We are experiencing far fewer support problems with critical business applications that have strict host-to-host latency requirements, such as our Intelligent Contact Management (ICM) system," adds Tom Wojciaczyk, Cisco IT network engineer.

In addition to lessened support requirements for real-time applications, Cisco has been able to more intelligently address bandwidth capacity issues. In the past, IT would respond to complaints that the network was slow and unresponsive by pushing for a capacity upgrade. Now, says Packer, "We prioritize applications through the hub. We may still run real-time and batch applications simultaneously, but we have the knowledge now to heavily throttle back on batch applications. QoS has given us real leeway."

Proper design and implementation of QoS has enabled Cisco to develop consistent standards across its hardware and software infrastructure. The QoS architecture provides a baseline for driving network upgrades—and is easier to deploy across various geographic regions.

A consistent QoS configuration has made support and troubleshooting simpler worldwide. "There are still operational challenges," says Nguyen, "but with some training, our global support teams can roll out new applications and do the appropriate management at the WAN and Layer-2 LAN edges with very few problems. They no longer need to mark traffic at every hop for its transit across the network, and they can confine most of their troubleshooting to where traffic enters and leaves the WAN and the LAN. "

LESSONS LEARNED

The Cisco IT QoS team can point to important lessons learned in the design and implementation of QoS V2 that can be applied to all major enterprise QoS deployments. Among them:

- A standard architecture is critical to a smooth end-to-end QoS deployment across an enterprise landscape. Without standards, implementation of QoS is difficult and costly—as are application, upgrade, and patch distribution.
- In some instances, it may be necessary to supplement bandwidth allocations for certain real-time applications.
- A clear understanding of traffic and user requirements, as well as overall network topology and individual site operation, is essential before the design process begins.
- Partnering with other business units and technology teams improves both near- and long-term results. For example, the design team worked with several internal Cisco organizations to identify QoS features that could be designed into various switch and router products
- Maintaining standardized hardware platforms—and establishing Cisco IOS® Software-based and Cisco Catalyst® Operating System Software-based versions—ensures that configurations can be simplified and automated.

NEXT STEPS

With the emergence of video telephony and other demanding real-time applications, the Cisco IT QoS design team is at work on enhancements to QoS V2 that will standardize and automate QoS configurations, making them easier to "push" and manage across the enterprise. The team is also building real-time monitoring capabilities designed to make troubleshooting easier, improve the quality of reports, and further enhance network capacity planning. And efforts are underway with the internal enterprise management developers and the Network Management Technology Group Business Unit to establish a pilot program with IT to enhance Cisco QoS Policy Manager (QPM).

FOR MORE INFORMATION

To read the entire case study or for additional Cisco IT case studies on a variety of business solutions, visit Cisco on Cisco: Inside Cisco IT www.cisco.com/go/ciscoit

NOTE

This publication describes how Cisco has benefited from the deployment of its own products. Many factors may have contributed to the results and benefits described; Cisco does not guarantee comparable results elsewhere.

CISCO PROVIDES THIS PUBLICATION AS IS WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Some jurisdictions do not allow disclaimer of express or implied warranties, therefore this disclaimer may not apply to you.

Americas 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: 408 527-0883

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

Europe Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: +31 0 800 020 0791
Fax: +31 0 20 357 1100

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**