

IP QoS FAQ

Frequently Asked Questions about IP Quality of Service



Table of Contents

Section 0. Introduction

- 0.1 Maintainer and Affiliation
- 0.2 What if the information you need is not in the FAQ?
- 0.3 How do you contribute to the FAQ?
- 0.4 Legal stuff
- 0.5 Acknowledgements

Section 1 - What is Quality of Service?

- 1.1 Definitions of Quality of Service
- 1.2 Why Quality of Service is needed
- 1.3 Quality of Service parameters
- 1.4 Quality of Service versus Class of Service (Cos)

Section 2 - What are the business benefits of QoS?

- 2.1 What are the benefits to applications?
- 2.2 What are the benefits to enterprises?
- 2.3 What are the benefits to service providers?

Section 3 - QoS Architecture Models for Traffic Engineering

- 3.1 Over Provision
- 3.2 Reservation Based - Integrated Services Architecture (Int-Serv)
- 3.3 Prioritization (reservation-less) - Differentiated Services Framework (Diff-Serv)

Section 4 - What are Key QoS Parameters?

- 4.1 Latency
- 4.2 Jitter
- 4.3 Bandwidth
- 4.4 Packet Loss
- 4.5 Availability

Section 5 - What are Service Level Agreements (SLA's)?

- 5.1 SLA Parameters
- 5.2 Service Level Specifications
- 5.3 Traffic Conditioning Specifications
- 5.4 Examples

Section 6 - QoS Policies

- 6.1 Policies
- 6.2 Mechanisms
- 6.3 Data Store (DEN, DTMF)
- 6.4 Policy Decision Points (PDP)
- 6.5 Policy Enforcement (PEP)
- 6.6 Policy Protocols(COPS)
- 6.7 Bandwidth Brokers

Section 7 - What is the Differentiated Services Framework (DiffServ)?

- 7.1 How Diff-Serv works
- 7.2 TOS
- 7.3 Per Hop Behavior (PHB) Policies (Forwarding)
- 7.4 Traffic classification
- 7.5 Traffic conditioning
- 7.6 Current status (IETF status, industry deployment)

Section 8 - What is the Resource Reservation Setup Protocol (RSVP) ?

- 8.1 Relationship to the Integrated Services Architecture (Int-Serv)
- 8.2 How RSVP works
- 8.3 Current status (IETF status, industry deployment)

Section 9 - What are the key QoS mechanisms?

- 9.1 Admission control
- 9.2 Traffic shaping/conditioning
- 9.3 Packet classification
- 9.4 Packet marking
- 9.5 Scheduling and priority mechanisms
- 9.6 Signalling protocols
- 9.7 Queuing (WFQ, CFQ, SFQ)
- 9.8 Congestion Control (RED, ECN)

Section 10 - What is MPLS?

- 10.1 MPLS overview
- 10.2 MPLS with Diff-Serv
- 10.3 MPLS with RSVP
- 10.4 MPLS with LDP
- 10.5 Current status (IETF status (stable or draft RFC), industry deployment)

Section 11 - What about 802.1p/Q?

- 11.1 802.1p
- 11.2 802.1Q
- 11.3 802.1D

Section 12 - What about QoS in Layer 2 and Layer 3 switching?

- 12.1 Layer 2
- 12.2 Layer 3

Section 13 - How does routing support QoS?

- 13.1 Diff-serv, Int-serv routing
- 13.2 Routing at ingress
- 13.3 Interior routing
- 13.4 Interdomain considerations

Section 14 - What are some implementation considerations?

- 14.1 Large Service providers and interdomain considerations
- 14.2 Within a customer network

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

14.3 Interoperability
14.4 QoS support in application software
14.5 Monitoring and measuring QoS
14.6 Managing QoS
14.7 QoS solutions for IP Multicast
14.8 Security and authentication
14.9 Billing, pricing and accounting
14.10 Testbeds (QBone)

Section 15 - IP QoS over other network infrastructures

15.1 ATM
15.2 IP QoS over ATM using layer 3 switching in carrier networks
15.3 Frame relay

Section 16 - Where can I find developer information?

16.1 Windows: Winsock2 and GQoS API

Section 17 - What's going on standards committees?

17.1 IETF
17.2 ITU

Section 18 - Useful References

18.1 Books
18.2 Mailing lists
18.3 Web sites
18.4 IETF

0. Introduction

0.1 Maintainer and affiliation

Please send contributions to the FAQ to the maintainer, Vicki Johnson (faq@stardust.com).
Maintainer's affiliation: The QoS Forum (www.qosforum.com), an initiative to promote the understanding and use of QoS. The QoS Forum is comprised of leading vendors of QoS products and services under the leadership of Stardust Forums (www.stardust.com).

0.2 What if the information you need is not in the FAQ?

(1) Review the Technology Central resources at the QoS Forum web site, www.qosforum.com.
(2) See the References section of this FAQ.
(3) Contact vendors of QoS products and services. See the QoS Buyer's Guide MarketPlace at www.qosforum.com for a vendor list.

0.3 How do you contribute to the FAQ?

Please send contributions to the FAQ to the maintainer, Vicki Johnson (faq@stardust.com).

0.4 Legal stuff

Copyright © 1999 Stardust Forums, Inc. All Rights Reserved. The text of this publication, or any part thereof, may not be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, storage in an information retrieval system, or otherwise, without prior written permission of Stardust Forums, Inc. Stardust is a registered trademark and Stardust Forums is a trademark of Stardust Technologies, Inc. The QoS Forum is a division of Stardust Forums, Inc.

Stardust Forums, Inc. does not itself distribute, ship or sell nor permit others to distribute, ship or sell its copyrighted materials to individuals or businesses in countries that are not members of the Berne Convention or the Universal Copyright Convention.

RESTRICTED RIGHTS LEGEND USE, DUPLICATION, OR DISCLOSURE BY THE GOVERNMENT IS SUBJECT TO RESTRICTIONS AS SET FORTH IN SUBPARAGRAPH (c) (1) (ii) OF THE RIGHTS IN TECHNICAL DATA AND COMPUTER SOFTWARE CLAUSE AT DFARS 252.227-7013 or subparagraphs (c) (1) and (2) of Commercial Computer Software—Restricted Rights at 48 CFR 52.227-19, as applicable.

There is no warranty of any kind, express or implied, for the information in this FAQ, including warranties of accuracy, suitability or fitness for a particular purpose. Using this information implies your acceptance of these terms.

0.5 Acknowledgements

Kai Chang Kai_Chang@3com.com, Vic Forgetta Vic_Forgetta@netcomsystems.com, Martin Hall martinh@stardust.com, Marjory Johnson mji@riacs.edu, Vicki Johnson vji@interconnect.com, Jae Lee jaelee@newbridge.com, Tony Lee tleee@extremenetworks.com, Trinh Ngo ngo@nexen.com, Elizabeth Racioppi Elizabeth_Racioppi@corpeast.BayNetworks.com, Steve Richman srichman@lucent.com, Bob Quinn rcq@stardust.com.

Section 1 - What is Quality of Service ?

1.1 Definitions of Quality of Service

In the simplest sense, Quality of Service (QoS) means providing consistent, predictable data delivery service. In other words, satisfying customer application requirements.

QoS is to the ability of a network element (e.g. an application, host or router) to have some level of assurance that its traffic and service requirements can be satisfied. To enable QoS requires the cooperation of all network layers from top-to-bottom, as well as

To join the Quality of Service, contact [Nancy Moss](mailto:Nancy_Moss) at (408)879-8080 or register with QoSF.

every network element from end-to-end. Any QoS assurances are only as good as the weakest link in the "chain" between sender and receiver.

QoS does not create bandwidth. It isn't possible for the network to give what it doesn't have, so bandwidth availability is a starting point. QoS only manages bandwidth according to application demands and network management settings, and in that regard it cannot provide certainty if it involves sharing. Hence, QoS with a guaranteed service level requires resource allocation to individual data streams. A priority for QoS designers has been to ensure that best-effort traffic is not starved after reservations are made. QoS-enabled (high-priority) applications must not disable the mundane (low-priority) Internet applications.

1.2 Why Quality of Service is needed

The Internet Protocol (IP), and the architecture of the Internet itself, is based on the simple concept that datagrams with source and destination addresses can traverse a network of (IP) routers independently, without the help of their sender or receiver. The Internet was historically built on the concept of a dumb network, with smarts at either end (at the sender and receiver).

There is a price to pay for this simplicity, however. The reason IP is simple is because it doesn't provide many services. IP provides addressing, and that enables the independence of each datagram. IP can fragment datagrams (in routers) and reassemble them (at the receiver), and that allows traversal of different network media. But IP does not provide reliable data delivery. Routers are allowed to discard IP datagrams en route, without notice to sender or receiver. IP relies on upper-level transports (e.g. TCP) to keep track of datagrams, and retransmit as necessary. And these "reliability" mechanisms can only assure data delivery; neither IP nor its high-level protocols can ensure timely delivery or provide any guarantees about data throughput. IP provides what is called a "best effort" service. It can make no guarantees about when data will arrive, or how much it can deliver.

This limitation has not been a problem for traditional Internet applications like web, email, file transfer, and the like. But the new breed of applications, including audio and video streaming, demand high data throughput capacity (bandwidth) and have low-latency requirements when used in two-way communications (i.e. conferencing and telephony). Public and private IP Networks are also being used increasingly for delivery of mission-critical information that cannot tolerate unpredictable losses.

Unlike "pure virtual circuit" technologies like ATM and Frame Relay, IP does not make hard allocations of resources. This provides much more efficient use of the available bandwidth, and it is more flexible also. Typical network traffic is bursty rather than continuous. IP is datagram-based so it uses the available bandwidth most efficiently, by sharing what is available as needed. This also allows IP to adapt more flexibly to applications with varying needs. However, it also leads to some unpredictability in service. The capacity to tolerate this unpredictability relates to the level of guarantee they require.

1.3 Quality of Service parameters

There are a number of characteristics that qualify QoS, including

- minimizing delivery delay
- minimizing delay variations
- providing consistent data throughput capacity

1.4 Quality of Service versus Class of Service (Cos)

The focus of QoS is providing predictable service (service defined as the share of available capacity) during periods of congestion. It is the periods of congestion that are the target of QoS. Being able to measure and report on service quality is also an important attribute of QoS solutions. Class of Service is a more general term for such capabilities.

Section 2 - What are the business benefits of QoS ?

2.1 What are the benefits to applications?

The Internet is increasingly relied upon for doing business, and the expectations for quality assurances are the same as for a private, controlled network. The Internet is being used to power both intranets within the enterprise and extranets that enable electronic commerce with business partners. As business is increasingly conducted over the web, it becomes more important that IT managers ensure that these networks deliver appropriate levels of quality. Quality of Service (QoS) technologies provide the tools for IT managers to deliver mission critical business over the public network.

2.2 What are the benefits to enterprises?

Applications are getting more demanding. Mission-critical applications deployed over IP networks increasingly require quality, reliability, and timeliness assurances. In particular, applications that use voice, video streams, or multi-media must be carefully managed within an IP network to preserve their integrity.

Managing QoS becomes increasingly difficult because many applications deliver unpredictable bursts of traffic. For example, usage patterns for web, email, and file transfer applications are virtually impossible to predict, yet network managers need to be able to support mission-critical applications even during peak periods.

QoS technologies allow IT managers and network managers to:

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

- Manage jitter sensitive applications, such as audio and video playbacks
- Manage delay-sensitive traffic, such as real time voice
- Control loss in times of inevitable bursty congestion

2.3 What are the benefits to service providers?

Clearly, enterprises and corporations are becoming more attuned to the requirements of mission-critical business conducted over the public network. Increasingly, they are also outsourcing more and more network services to service providers, which allows them to focus more on the internal business and reduce capital expenses. This means that service providers who can offer quality assurances for end-to-end business traffic will win more enterprise business going forward. QoS technologies will allow service providers to offer more services, such as real-time traffic support, or specific bandwidth allocations, to build into an SLA portfolio. This creates more revenue generation for service providers, while offering more services to enterprises.

Contributions to section 2 from [Nortel](#).

Section 3 - QoS Architecture Models for Traffic Engineering

3.1 Over Provision

The obvious solution to handle peak periods is to over-provision the network, to provide surplus bandwidth capacity in anticipation of these peak data rates during high-demand periods. Equally obvious, however, is that this is not economically viable--at least not with today's bandwidth technologies and infrastructures -- and especially for WAN links. Since peak data rates and the network regions on which they might occur are seldom possible to predict, this is not a realistic alternative anyway.

IP is necessary and bandwidth is necessary, but neither is sufficient for all application needs under all conditions. Best effort cannot always provide a usable service, let alone an acceptable one. Even on a relatively unloaded IP network, delivery delays can vary enough to adversely affect applications that have real-time constraints. To provide service guarantees--some level of quantifiable reliability--IP services must be supplemented with the ability to differentiate traffic and enable different service levels for different users and applications. The following two types of QoS available are complementary, designed for use in combination for different network contexts.

3.2 Reservation Based - Integrated Services Architecture (IntServ)

The Integrated Services Architecture being defined by the IETF is intended to transition the Internet into a robust integrated-service communications infrastructure that can support the transport of audio, video, real-time, and classical data traffic. Network resources are apportioned according to an application's QoS request, and subject to bandwidth management policy. RSVP ([see Section 8 RSVP](#)) provides the mechanisms to do this, as a part of the IntServ architecture.

3.3 Prioritization (reservation-less) - Differentiated Services Framework (DiffServ)

The Differentiated Services Framework being defined by the IETF is intended to meet the need for relatively simple and coarse methods of providing differentiated classes of service for Internet traffic, to support various types of applications, and specific business requirements. The differentiated services approach to providing quality of service in networks employs a small, well-defined set of building blocks from which a variety of services may be built. A small bit-pattern in each packet, in the IPv4 TOS octet or the IPv6 Traffic Class octet, is used to mark a packet to receive a particular forwarding treatment, or per-hop behavior, at each network node. Network traffic is classified and apportioned network resources according to bandwidth management policy criteria. To enable QoS, classifications give preferential treatment to applications identified as having more demanding requirements. DiffServ ([see Section 7 DiffServ](#)) provides this service.

Section 4 - What are Key QoS Parameters?

4.1 Latency

The time between a node sending a message and receipt of the message by another node. This encompasses delay in a transmission path or in a device within a transmission path. In a router, latency is the amount of time between when a data packet is received and when it is retransmitted. Also referred to as propagation delay.

4.2 Jitter

An aberration that occurs when video or voice is transmitted over a network, and packets do not arrive at its destination in consecutive order or on a timely basis, i.e. they vary in latency. The distortion of a signal as it is propagated through the network, where the signal varies from its original reference timing. In packet-switched networks, jitter is a distortion of the interpacket arrival times compared to the interpacket times of the original transmission. This distortion is particularly damaging to multimedia traffic. For example, the playback of audio or video data may have a jittery or shaky quality.

4.3 Bandwidth

A measure of data transmission capacity, usually expressed in kilobits per second (Kbps) or megabits per second (Mbps). Bandwidth indicates the theoretical maximum capacity of a connection, but as the theoretical bandwidth is approached, negative factors such as transmission delay can cause deterioration in quality. If you increase bandwidth, you can transfer more data. Network bandwidth can be visualized as a pipe that transfers data. The larger the pipe, the more data can be sent through it.

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

4.4 Packet Loss

Example: 1% or less on network-wide monthly average packet loss.

4.5 Availability

Example: 99.9% premises to POP.

Section 5 - What are Service Level Agreements (SLA's)?

5.1 SLA Parameters

A Service Level Agreement (SLA) is a service contract between Service provider and their customer that defines provider responsibilities in terms of network levels (throughput, loss rate, delays and jitter) and times of availability, method of measurement, consequences if service levels aren't met or the defined traffic levels are exceeded by the customer, and all costs involved. It specifies the forwarding service a customer should receive. A SLA may include traffic conditioning rules.

5.2 Service Level Specifications

The details of the operational characteristics for an SLA are further defined in terms of Service Level Specifications (SLS) and/or Objectives (SLOs). The SLS may consist of expected throughput, drop probability, latency, constraints on the ingress and egress points at which the service is provided, indicating the `scope' of the service, traffic profiles which must be adhered to for the requested service to be provided, disposition of traffic submitted in excess of the specified profile, and marking and shaping services provided.

An SLO partitions an SLA into individual objectives that can be mapped into policies that can be executed. The SLOs define metrics to enforce, police, and/or monitor the SLA. Some commonly used metrics to determine whether or not an SLA is being fulfilled include component system availability (e.g., up-time and MTBF), performance (e.g., response time), and serviceability (e.g., MTTR).

5.3 Traffic Conditioning Specifications

Traffic conditioning control functions that can be applied to a behavior aggregate, application flow, or other operationally useful subset of traffic, e.g., routing updates. These may include metering, policing, shaping, and packet marking. Traffic conditioning is used to enforce agreements between DiffServ domains, for example.

A Traffic Conditioning Agreement (TCA) is an agreement specifying classifier rules and any corresponding traffic profiles and metering, marking, discarding and/or shaping rules which are to apply to the traffic streams selected by the classifier. A TCA encompasses all of the traffic conditioning rules explicitly specified within a SLA along with all of the rules implicit from the relevant service requirements.

5.4 Examples

IP Services Metrics	Service Level
Network Availability	99.9% premises to POP
Latency	80 ms or less network-wide monthly average delay
Latency	Traffic offered at service level A will be delivered with low latency
Packet Loss	1% or less on network-wide monthly average packet loss
Packet Loss	Traffic offered at service level B will be delivered with low loss
Throughput	99.99% network-wide data delivery rate average per calendar month

Section 6 - QoS Policies

6.1 Policies

In an open and public Internet (as well as large intranets), the acceptance of QoS requests results in better network service to some flows, possibly at the expense of service to traditional best-effort flows. The purpose of such classification may include preferential queuing or dropping, admitting or denying access, or encrypting the packet's payload, to cite just a few examples. Protocols that explicitly support some or all of these functions include COPS, RADIUS, RSVP, IntServ, DiffServ, ISSLL, DSSLL, and IPSEC. The successful wide-scale deployment these and other protocols depends on the ability for the administrator of a network domain to administer and distribute consistent policy information to the multiple devices in the network which perform the classification and packet conditioning or treatment. Protocols that could be used for the distribution of the policy include LDAP, COPS, SNMP, and TELNET/CLI. The multiple types of devices that must work in concert across even a single domain to achieve

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

the desired policy can include hosts (clients and servers), routers, switches, firewalls, bandwidth brokers, subnet bandwidth managers, network access servers, and policy servers, to name just a few.

The IETF Policy working group is working on how to represent, manage, and share policies and policy information in a vendor-independent, interoperable, scalable manner for QoS traffic management. See the working group page at <http://www.ietf.org/html.charters/policy-charter.html>.

Examples of internet-drafts from this working group are [Policy Action Classes for Differentiated Services and Integrated Services](#), [Terminology for describing network policy and services](#), [Policy Framework Core Information Model](#), and [Policy Framework Definition Language](#). No RFC's have been published at this time.

Some of the terminology and concepts below are drawn from the above documents.

6.2 Mechanisms

Policy is comprised of the following three functions:

- Decision-making. This compares the current state of the network to a desired state described by an application-specific policy and decides how to achieve the desired state.
- Enforcement. This implements a desired policy state through a set of management commands; when applied to network elements, these management commands change the configuration of the device using one or more mechanisms. These mechanisms MAY be vendor-specific.
- Policing. This is an on-going active or passive examination of the network and its constituent devices for checking network health, whether policies are being satisfied, and whether clients are taking unfair advantage of network services. Decision-making uses static and/or dynamic data to determine if a type of policy is being satisfied and, if not, what steps are needed to satisfy that policy. Enforcement refers to the interpretation and execution of policies by consumers who are assumed to be trustworthy. Policing is the auditing of policy compliance to verify that the policy consumers properly implement policies.

A general policy architecture is shown below.

Example: Provide the JitterFreeMPEG2 video service for authorized users between authorized points, but only at agreed-upon times. The policy condition could be loosely translated as: IF the user is a member of an approved group (ApprovedUsers) that are authorized to have this service) AND the service requested is one supported (VideoServicesgroup) AND the source of the request is approved (in the VideoSources group or has been authenticated) AND the destination is approved (in the VideoDestinations group or has been authenticated) AND the time requested is OK (in ApprovedTimePeriods) .

6.3 Data Store (DEN, DTMF)

This repository may be, but is not limited to, a directory accessed using the LDAP protocol. The Directory Enabled Network (DEN) Initiative and related specification work is an effort to build intelligent networks and networked applications that can associate users and applications to services available from the network according to a consistent and rational set of policies. DEN defines a directory as a centralized repository that coordinates information storage and retrieval, enabling other data- and application-specific repositories to be united. Eventually, intelligent network applications will transparently leverage the appropriate information about the network and the services that it offers on behalf of its users and particular context that the application is running in. See the DEN FAQ at the Desktop Management Forum site www.dmtf.org.

6.4 Policy Decision Points (PDP)

The point where policy decisions are made.

6.5 Policy Enforcement (PEP)

The point where the policy decisions are actually enforced. It is assumed that policy decisions will always be made in the PDP and implemented in the PEP. Specifically, the PEP is not able to make decisions on its own. This simplifies the definition and modeling of policy.

6.6 Policy Protocols (COPS)

The Common Open Policy Service (COPS) protocol is emerging as a viable solution for distributed policy management. Initially, COPS will be used within a domain, for router policy enforcement points (PEPs) to retrieve policy from the policy distribution points (PDPs). COPS may be also be used between Bandwidth Brokers (BBs)-which essentially act as PDPs-for dynamic inter-domain policy exchange. Bandwidth Brokers could be third parties that manage SLAs for various ISPs and enterprises. Border exchanges of data between administrative domains can be policed, shaped and conditioned according to SLAs that are encoded in a generic policy grammar.

Examples of internet-drafts being produced by the RSVP Admission Policy (rap) IETF working group are [A Framework for Policy-based Admission Control](#), [The COPS \(Common Open Policy Service\) Protocol](#), [RSVP Extensions for Policy Control](#) and [COPS usage for RSVP](#) .

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

See the IETF RAP working group page <http://www.ietf.org/html.charters/rap-charter.html>.

6.7 Bandwidth Brokers

Bandwidth brokers for handling QoS policies are being defined in association with DiffServ. The following explanation is drawn from <http://ietf.org/internet-drafts/draft-nichols-diff-svc-arch-01.txt>

Bandwidth brokers (BB) are designed to be configured with organizational policies, keep track of the current allocation of marked traffic, and interpret new requests to mark traffic in light of the policies and current allocation. They are intended to be used to allocate bandwidth for end-to-end connections with less state and simpler trust relationships than deploying per flow or per filter guarantees in all network elements on an end-to-end path. Organizationally, the BB architecture is motivated by the observation that multilateral agreements rarely work and this architecture allows end-to-end services to be constructed out of purely bilateral agreements. BBs only need to establish relationships of limited trust with their peers in adjacent DiffServ domains, unlike schemes that require the setting of flow specifications in routers throughout an end-to-end path. BBs have two responsibilities. Their primary one is to parcel out their region's Marked traffic allocations and set up the leaf routers within the local domain. The other is to manage the messages that are sent across boundaries to adjacent regions' BBs. A BB is associated with a particular trust region, one per domain. A BB has a policy database that keeps the information on who can do what when and a method of using that database to authenticate requesters. Only a BB can configure the leaf routers to deliver a particular service to flows, crucial for deploying a secure system. An initial request might cause communication between BBs on several domains along a path, but each communication is only between two adjacent BBs. Initially, these agreements will be prenegotiated and fairly static. Some may become more dynamic as the service evolves.

Section 7 - What is the Differentiated Services Framework (DiffServ)?

Differentiated Services (DiffServ) is an IETF specified QoS mechanism that handles traffic flows in one or more networks. DiffServ supplies a very customizable QoS, depending on the needs of the traffic type. DiffServ calls for pushing major state and work to edges of network, while forwarding is expected to be done very quickly in the core of the network. In the DiffServ framework, packets carry their own state in a few bits of the IP header (the DS CodePoint), which also leads to scalability of this QoS mechanism, making it appropriate for end-to-end QoS. Policy decisions and implementations are left to local "trust" domains. The IETF is working to standardize DiffServ behaviors.

DiffServ was proposed to provide service differentiation by creating service classes with different priorities using either the type-of-service (TOS) field of IPv4 or the priority bits of IPv6 headers. This priority scheme translates into higher throughput for higher priority classes. The goal for differentiated services is to define a scalable service discrimination without the need for per-flow state and signaling at every hop.

Network applications have performance and consistency expectations at some level. Treating all traffic at the same priority, leads to major difficulties when bandwidth is restricted. In order to create a single multiservice network that caters to various applications with different characteristics and priorities, methods must be found to ensure that the network performance is acceptable for each application. Critical to all these improvements is a fundamental shift from a single priority networking to having service differentiation for different traffic flows.

DiffServ, defined in RFC 2475 and 2474, defines a scalable service discrimination policy without maintaining the state of each flow and signalling at every hop. The primary goal of differentiated services is to allow different classes of service to be provided for traffic streams on a common network infrastructure. Differentiated Service aggregates multitude of QoS-enabled flows into a small number of aggregates. These aggregated flows are given differentiated treatment within the network. The diffserv approach attempts to push per-flow complexity away from the network core and towards the edge of the network where both the forwarding speeds and the fan-in of flows are smaller. Each diffserv flow is policed and marked according to the service profile at the edge of the network, and service only a small number of traffic aggregates in the core.

DiffServ will provide a controlled and coarsely predictable IP class of service. To support different classes of IP service over the internet, the IP differentiated services architecture defines three main building blocks: packet classifiers, forwarding/per-hop-behavior, and traffic conditioning policies. The differentiated services model utilizes static configuration of classification and forwarding policies in each node along a network path.

7.1 How Diff-Serv works

Each packet receives a particular forwarding treatment based on marking in its IP TOS octet (now called DS CodePoint). The packet may be marked anywhere in the network, but probably at domain boundaries. The packet is treated the same way as others marked the same. There is no per-flow state required inside the network; core devices know only markings, not flows. Per-flow state kept at the network edge, that is, flows are aggregated based on desired behavior. A note - this scheme requires overall network engineering so that aggregates get the appropriate or desired services.

Services are built by applying rules: Rules for how packets are marked initially; rules for how marked packets are treated at boundaries. At boundaries of domains, the only requirement is to have bilateral agreement between the parties on each side of the boundary (i.e., no multilateral agreements required). Three elements work together to deliver a Diffserv service:

- Per-Hop Behaviors (PHBs) - deliver special treatment to packets at forwarding time
- Traffic Conditioners - alter packet aggregates to enforces rules for services
- Bandwidth Brokers (Policy Managers) - apply and communicate policy

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

The IETF DiffServ group has defined two service classes for supporting applications. The Premium Service model emulates traditional leased line service that promises to deliver customer traffic with a low loss probability at a given peak rate, and a strict priority based per-hop behavior for ensuring low latency. This service is appropriate to applications that have strict bandwidth, latency, and jitter requirements. To create a low-loss, low-delay service, nodes must be configured such that the aggregate has a well-defined minimum departure rate, independent of other traffic. The second service model, Assured Service model emulates a lightly loaded network even in the presence of congestion. This service promises to deliver traffic with a high degree of reliability within the negotiated latency limits. The Assured Forwarding PHB group provides delivery of IP packets in four independently forwarded classes with three-discard precedence values for each class.

7.2 TOS

The ToS field, now called the DiffServ codepoint (DSCP), is an 8-bit field, the last 2 bits are reserved. With the 6-bit DSCP, there are 64 possible codepoints; 48 in the global space, 16 for local use. Host vendors want to be able to set DSCP, and may also want to check and possibly reset DSCP at domain boundaries.

7.3 Per Hop Behavior (PHB) Policies (Forwarding)

PHBs are the packet forwarding treatment that delivers the "differentiated service" to packets at the network node output: policing, shaping, possible remarking of DS Code Point, enqueueing treatment (e.g., drop preference), scheduling.

Currently the IETF is defining the following PHBs: Expedited Forwarding (EF), Assured Forwarding (AF), and Default (DE).

Expedited Forwarding requires that (at each node) the egress rate exceed the ingress rate for a conforming aggregate. This is like a "Virtual leased line." The EF treatment polices and drops on network ingress and shapes on egress to maintain the service contract to next provider. Modest buffering is needed (no burst) and some form of priority queueing is required. This treatment should be kept to a small fraction of total network traffic. This is a proposed IETF Standard.

Assured Forwarding defines 4 priorities of traffic receiving different bandwidth levels (the "olympic services" Gold - Silver - Bronze - Best Effort). There are 3 drop preferences each (similar to Frame Relay in this respect). The worse the drop preference, the more chance of getting dropped during congestion. There are token bucket policers for each priority (burst allowed). Enqueueing uses RED or similar mechanism to distinguish drop preference and control congestion and scheduling based on the bandwidth of the priority.

The DS byte restructures the TOS field in the IPv4 header to permit use of parameters relevant to specified service levels and traffic behavior controls. Of the 8 bits in the field, 6 bits define the per-hop behavior (PHB) the packet will receive with respect to policies established at a network boundary; and the rest of the 2 bits are *currently unused* (CU). The following figure shows the Ipv4 datagram with the DS type field included.

At the edge of the network or administrative boundary, the classifier determines the value of the DS field for each incoming flow. Based on the DS field, these incoming flows are aggregated over an outgoing flow. The Router implementations in the intermediate nodes use this 6-bit PHB field to index into a table for selecting a particular packet-handling mechanism. This forwarding policy determines how routers will handle the packets in terms of providing a class of service by combining traffic management functions, such as packet queueing, scheduling, and buffer reservations at each node.

Diffserv expects advance provisioning and reservations made in each of the intermediate nodes along the network path. If a network path crosses multiple DS domains or multiple ISPs, the ISPs must support the same PHBs to provide a consistent end-to-end service.

7.4 Traffic classification

The classifier selects packets based on the combination of one or more predefined set of header fields. The mapping of network traffic to the specific behaviors that result in different class of service is indicated by the Differentiated Service (DS) field shown below. Each DS field uniquely identifies the per-hop-behavior or the treatment given to the traffic at each hop along the network path. The diffserv architecture supports a maximum of 64 classes of service. Each router sorts the packets into queues based on the DS field. The queues might get different treatment based on their priority, share of bandwidth, and discard policies.

7.5 Traffic conditioning

Traffic Conditioners enforce the rules of each service at the network node input. Classifiers: both DSCP (TOS octet for IPv4) and general packet header ("deep classification"). Policers: "token bucket" with various actions Shapers: lets "customer" make aggregate flows conform Markers: set DS byte based on classification.

The Differentiated Services architecture offers a framework within which service providers can offer each customer a range of network services that are differentiated on the basis of performance in addition to pricing of tiers used in the past. These services are monitored for fairness and in meeting the service agreements. In order to deliver service agreements, each DiffServ enabled edge router implements Traffic Conditioning function which performs metering, shaping, policing and marking of packets to ensure that the traffic entering a differentiated services network conforms to the Traffic Conditioning Agreement (TCA).

This agreement encompasses all of the traffic conditioning rules specified within a Service Level Agreement.

- Metering - Monitors the traffic pattern of each flow against the traffic profile. For out-of-profile traffic the metering function interacts with other components to either re-mark, or drops the traffic for that flow.

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

- Marking - Customers request a specific performance level on a packet by packet basis, by marking the DS field of each packet with a specific value. This value specifies the Per-hop Behavior (PHB) to be allocated to the packet within the provider's network. The edge routers classify the packets to identify the PHB and a DS code point for that packet.
- Policing - At the ingress edge routers, the incoming traffic is classified into aggregates. These aggregates are policed according to the TCA. The out-of-profile traffic is either dropped at the edge or is remarked with a different PHB.
- Shaping - The routers control the forwarding rate of packets so that flow does not exceed the traffic rate specified by its profile. The shapers ensure fairness between flows that maps to the same class of service, and controls the traffic flow to avoid congestion.

The following figure shows an example of a network using DiffServ.

7.6 Current status (IETF status (stable or draft RFC), industry deployment)

IETF RFC's are [Definition of the Differentiated Services Field \(DS Field\) in the IPv4 and IPv6 Headers \(RFC 2474\)](#) and [An Architecture for Differentiated Services \(RFC 2475\)](#). See the IETF Diffserv working group page at <http://www.ietf.org/html.charters/diffserv-charter.html>.

DiffServ products and services are expected to be offered to customers by beginning of 2000.

Contributions to section 7 from [Nortel](#), [Fujitsu Nexion](#).

Section 8 - What is the Resource Reservation Setup Protocol (RSVP)?

To accommodate a diverse set of applications and to enrich the best-effort service model, the Internet Engineering Task Force (IETF) is considering a number of extensions that permit the allocation of different levels of service to different users. One of the outcomes of the effort is Resource ReSerVation Protocol (RSVP), a signaling protocol for resource reservation. RSVP can be used to offer service discrimination for delay sensitive applications by explicit allocation of resources in the network. Its major features include: (1) the use of "soft state" in the routers, (2) receiver-controlled reservation requests, (3) flexible control over sharing of reservations and forwarding of subflows, and (4) the use of IP multicast for data distribution.

8.1 Relationship to the Integrated Services Architecture (IntServ)

Work is underway in the IETF to expand the Internet service model so that its packet switching protocols can support Integrated Services (IntServ) --- the transport of audio, video, real-time, and classical data traffic -- within a single network infrastructure. The IntServ framework provides the ability for applications to choose among multiple, controlled levels of delivery service for their data packets. To support this capability, two things are required: (1) individual network elements (subnets and IP routers) along the path followed by an application's data packets must support mechanisms to control the quality of service delivered to those packets. (2) a way to communicate the application's requirements to network elements along the path and to convey QoS management information between network elements and the application must be provided. In the integrated services framework the first function is provided by QoS control services such as Controlled-Load [RFC 2211] and Guaranteed [RFC 2212]. The second function may be provided in a number of ways, but is frequently implemented by a resource reservation setup protocol such as RSVP [RFC 2205].

The IntServ working group is defining the interfaces that express the application's end-to-end requirements, router scheduling interfaces that define what information is made available to individual routers within the network, and (general) subnet interfaces interfaces.

IntServ RFC's include [The Use of RSVP with IETF Integrated Services \(RFC 2210\)](#), [Integrated Services Management Information Base using SMIv2 \(RFC 2213\)](#), [Integrated Services Management Information Base Guaranteed Service Extensions using SMIv2 \(RFC 2214\)](#), [General Characterization Parameters for Integrated Service Network Elements \(RFC 2215\)](#), [Network Element Service Specification Template \(RFC 2216\)](#), [Specification of the Controlled-Load Network Element Service \(RFC 2211\)](#), [Specification of Guaranteed Quality of Service \(RFC 2212\)](#).

See the IETF IntServ page <http://www.ietf.org/html.charters/intserv-charter.html>.

8.2 How RSVP works

RSVP is a reservation setup and control protocol that provides a type of circuit-emulation on IP networks. RSVP is the most complex of all the QoS technologies, for applications (hosts) and for network elements (routers and switches). As a result, it also represents the biggest departure from the tried-and-true "best-effort" IP network, which creates some cause for concern.

Here is a simplified overview of how the protocol works:

- Senders characterize outgoing traffic in terms of the upper and lower bounds of bandwidth, delay, and jitter. RSVP sends PATH messages from the sender that contains this traffic specification (TSpec) information to the (unicast or multicast receiver(s)) destination address. Each RSVP-enabled router along the downstream route establishes a "path-state" that includes the previous source address of the PATH message (i.e. the next hop "upstream" towards the sender).

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

- To make a resource reservation, receivers send a RESV (reservation request) message "upstream" to the (local) source of the PATH message. In addition to the TSpec, the RESV message includes the QoS level required ([controlled] load or [guaranteed]) in an RSpec, and characterizes the packets for which the reservation is being made (e.g. the transport protocol and port number), called the "filter spec." Together, the RSpec and filter-spec represent "flow-descriptor" that routers use to identify reservations.
- When an RSVP router receives an RESV message, it uses the admission control process to authenticate the request and allocate the necessary resources. If the request cannot be satisfied (due to lack of resources or authorization failure), the router returns an error back to the receiver. If accepted, the router sends the RESV upstream to the next router.
- When the last router receives the RESV and accepts the request, it sends a confirmation message back to the receiver (note: the "last router" is either closest to the sender or at a reservation merge point for multicast flows).
- There is an explicit tear-down process for a reservation when sender or receiver ends an RSVP session.

The following figure shows RSVP "PATH" and "RESV" messages are used to establish a resource reservation between a sender and receiver. There is an explicit tear-down of reservations also (not shown).

Here are some salient characteristics of RSVP support:

- Reservations in each router are "soft," which means they need to be refreshed periodically by the receiver(s).
- RSVP is not a transport, but a network (control) protocol. As such, it does not carry data, but works in parallel with TCP or UDP data "flows."
- Applications require APIs to specify the flow requirements, initiate the reservation request, and receive notification of reservation success or failure after the initial request and throughout a session. To be useful, these APIs also need to include RSVP error information to describe a failure.
- Multicast reservations are "merged" at traffic replication points on their way upstream, which involves complex algorithms that are not well understood yet.
- Although RSVP traffic can traverse non-RSVP routers, this creates a "weak-link" in the QoS chain where the service falls-back to "best effort" (i.e. there is no resource allocation across these links).

8.3 Current status (IETF status, industry deployment)

The RSVP functional specification is described in the standards track [Resource ReSerVation Protocol \(RSVP\) -- Version 1 Functional Specification \(RFC 2205\)](#). Related RFC's are [RSVP Management Information Base using SMIv2 \(RFC 2206\)](#), [RSVP Extensions for IPSEC Data Flows \(RFC 2207\)](#), [Resource ReSerVation Protocol \(RSVP\) Version 1 Applicability Statement Some Guidelines on Deployment \(RFC 2208\)](#), and [Resource ReSerVation Protocol \(RSVP\) -- Version 1 Message Processing Rules \(RFC 2209\)](#). A number of IETF drafts have been published. See the IETF RSVP page at <http://www.ietf.org/html.charters/rsvp-charter.html>.

RSVP implementations are available in numerous vendor products. See the Buyer's Guide Marketplace at www.ipmulticast.com and www.qosforum.com.

Section 9 - What are the key QoS mechanisms?

9.1 Admission control

Admission Control determines whether a requested "connection" is allowed to be carried by the network. The main considerations behind this decision are current traffic load, current QoS, requested traffic profile, requested QoS, pricing and other policy considerations. For QoS enabled IP networks, Admission Control, for example, could be performed in the setting up of RSVP flows or MPLS paths.

9.2 Traffic shaping/conditioning

In QoS enabled IP networks, it's necessary to specify the traffic profile for a "connection" to decide how to allocate various network resources. Traffic Shaping/Conditioning ensures that traffic entering at an edge or a core node adheres to the profile specified. Typically, this mechanism is used to reduce the burstiness of a traffic stream. This involves a key tradeoff between benefits of shaping (e.g., loss in downstream network) and the shaping delay. Leaky Bucket based traffic shaping is an example of this mechanism.

9.3 Packet classification

In order to provide the requested QoS, it's critical to classify packets to enable different QoS treatment. This can be done based on various fields in IP headers (e.g., source/destination addresses and protocol type) and higher layer protocol headers (e.g., source/destination port numbers for TCP or UDP). Efficient and consistent Packet Classification is a key problem under active research.

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

9.4 Packet marking

Either as a result of a traffic monitoring mechanism or voluntary discrimination, a packet can be annotated for a particular QoS treatment in the network (e.g., high/low loss/delay priority). IP Packet Marking is proposed to be done using the IP header's Type of Service (TOS) byte for IPv4 and Traffic Class byte for IPv6.

9.5 Priority and scheduling mechanisms

To satisfy the QoS needs of different "connections," nodes need to have Priority and Scheduling Mechanisms.

The Priority feature typically refers to the capability of providing different delay treatment, e.g., higher priority packets are always served before the lower priority ones, both in the context of packet processing and transmission on outbound links. Nodes also implement different loss priority treatment, i.e., higher loss priority packets are lost less often than the lower loss priority ones.

Nodes also need to have the closely related Scheduling Mechanisms to ensure that different "connections" obtain their promised share of the resources (i.e., processing and link bandwidth). This mechanism also ensures that any spare capacity is distributed in a fair manner. Examples of this mechanism include Generalized Processor Sharing (GPS), Weighted Round Robin (WRR), Weighted Fair Queueing (WFQ), and Class Based Queueing (CBQ). Efficient implementation of these mechanisms, and extending them to include (a) both delay and bandwidth needs simultaneously, and (b) hierarchical scheduling are the areas of active research.

9.6 Signalling protocols

To obtain the required QoS from a network, end-systems need to signal the network the desired QoS as well as the anticipated offered traffic profile. This has been a fundamental part of various connection-oriented networks (e.g., ATM). However, for connectionless networks (e.g., IP), this is relatively new. Corresponding examples are the signaling associated with Resource ReSerVation Protocol (RSVP) and Label Distribution Protocol (LDP). Implementation scalability and the corresponding capabilities to signal different QoS needs are issues under current examination.

9.7 Queuing (WFQ, CFQ, SFQ)

Some network elements enable "fair queuing" algorithms so a misbehaving application--one that continues to send during times of congestion--won't punish other, better-behaved applications (e.g. TCP applications), or so the average of dropped packets is evenly distributed across flows [Queuing]. Basically, they determine how packets are dropped when congestion occurs in a router (i.e. when a queue is full). CFQ (Class-based Fair Queueing), WFQ (Weighted Fair Queueing), SFQ (Stochastic Fair Queueing) are examples of these algorithms.

9.8 Congestion Control (RED, ECN)

For QoS IP networks to operate in a stable and efficient fashion, it's essential that they have viable and robust Congestion Control capabilities. These capabilities refer to the ability to flow control and shed excessive traffic during the periods of congestion. Random Early Detection (RED) and Explicit Congestion Notification (ECN) are two of the proposed capabilities. RED prescribes discard probability to drop packets in a fair and robust way (i.e., consistent with behavior of higher layer protocols like TCP) based on the measured average queue length. RED (Random Early Detection) attempts to avoid congestion rather than reacting to it (and thereby avoid TCP synchronization problems that can result when hosts decrease or increase TCP traffic simultaneously after congestion occurs). It randomly drops packets before queues fill, to keep them from overflowing. Unlike the queue management algorithms mentioned above, it does not require flow-state in the routers.

ECN is a recently proposed mechanism for routers to notify existence of congestion to ECN-capable end-systems.

Contributions to section 9 by [Lucent](#).

Section 10 - What is MPLS?

10.1 MPLS overview

The rapid growth of the internet with ever-increasing number of users and traffic volume together with new real-time and multimedia applications, have created a need to improve forwarding technology in terms of bandwidth, and performance.

The MPLS working group within IETF is standardizing the technology for using a label-based forwarding paradigm in conjunction with layer-3 routing. The MPLS technology can operate over various link-level technologies, which includes packet-over-Sonet, frame relay, ATM, Ethernet, and token ring. MPLS combines layer 2 switching technology with layer 3 network layer services while reducing the complexity and operational costs.

Simplified Forwarding: Conventional routers using connectionless network layer protocols forward packets from one router to the next until the packet reaches its final destination. Each router along the routed path makes an independent forwarding decision by analyzing the packet header. The choice of next hop for a packet is based on the header analysis and result of running a routing algorithm. This approach is turning out to be insufficient to support today's networking demands, as routers are becoming a bottleneck.

MPLS defines a potential solution by taking a different approach to improve and simplify the packet forwarding function, and to provide sufficient network guarantees to support desired quality of service. To achieve this goal, MPLS adds connection-oriented mechanisms to connectionless network layer protocols. MPLS also adds new tools into connectionless network protocols for traffic

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

engineering and traffic management. The connection-oriented mechanisms identify pre-determined paths through the network between any two end-points. Each path is then assigned a label known as Label Switched Path (LSP).

At each ingress point of the network, an edge router known as Label Edge Router (LER) examines the IP header to determine the LSP. The LER then encapsulates the packet with MPLS header containing the label information, and the packet is forwarded to the next hop. All subsequent routers known as Label Switch Routers (LSR) use the label information from the header to determine the outgoing link and the new label for that outgoing link. The router then swaps the label in the MPLS header with the new label, and forwards the packet.

Label Distribution between Routers: Label-switched paths are controlled in a distributed fashion. Each router negotiates a label for each forwarding equivalence class (FEC) with its neighbors along the path. The forwarding equivalence class is derived using the incoming IP packet. Information on the topology of the network is maintained by one or more routing protocols such as OSPF, RIP, and BGP. For each route or aggregation of routes, a neighbor assigns a label. This information is distributed to the neighboring routers using Label Distribution Protocol (LDP). For each forwarding equivalence class, MPLS router maintains a mapping between an incoming label and interface, to an outgoing label and interface. These associations are stored in a database called Label Information Base (LIB).

Inter-Domain and Intra-Domain Forwarding with MPLS: IP routing views network as a set of administrative domains. To recognize domain boundaries, MPLS architecture defines a mechanism to allow label-switched paths to cross one or more domains. Inter-domain Label Switched Paths between two border routers are tunneled through an Intra-domain Label Switched Path. The MPLS packet transiting through the tunnel contains a stack of two labels for inter-domain and intra-domain switching. This mechanism helps to improve overall traffic performance and decreases the convergence time for routing.

10.2 MPLS with Diff-Serv

MPLS specifies ways to map Layer 3 traffic to a connection-oriented Layer 2 transports like ATM and frame relay. It adds a label containing specific routing information and allows routers to assign an explicit path to various classes of traffic. The MPLS network provides a transit service to DiffServ traffic by mapping DS-field to an explicit path through the MPLS network. For each Service Class, MPLS network provides an independent data flow path between the edge routers. The edge router obtains QoS information from the DS-field contained in the packet arriving from a DiffServ domain, and this information is used to map to an explicit path through the MPLS network.

10.3 MPLS with RSVP

IETF working group is currently investigating the possibility of using RSVP as a signaling protocol for establishing label switched paths, and propagating characteristics about the switched paths (e.g., Resource Reservation, Resource Re-assignment, QoS information, re-route information of label switched paths, loop detection, etc...) in MPLS networks.

10.4 MPLS with LDP

A fundamental concept in MPLS is that two Label Switching Routers (LSRs) must agree on the meaning of the labels used to forward traffic between and through them. This common understanding is achieved by using the Label Distribution Protocol (LDP). LDP is the set of procedures and messages by which Label Switched Routers (LSRs) establish Label Switched Paths (LSPs) through a network by mapping network-layer routing information directly to data-link layer switched paths. These LSPs may have an endpoint at a directly attached neighbor (comparable to IP hop-by-hop forwarding), or may have an endpoint at a network egress node, enabling switching via all intermediary nodes. LDP associates a Forwarding Equivalence Class (FEC) with each LSP it creates. The FEC associated with an LSP specifies which packets are "mapped" to that LSP. LSPs are extended through a network as each LSR "splices" incoming labels for a FEC to the outgoing label assigned to the next hop for the given FEC. See <http://ietf.org/internet-drafts/draft-ietf-mpls-ldp-03.txt>.

10.5 Current status (IETF status (stable or draft RFC), industry deployment

MPLS is an emerging standard from IETF. See the IETF MPLS working group page at <http://www.ietf.org/html.charters/MPLS-charter.html>.

Internet Drafts include [A Framework for Multiprotocol Label Switching](#), [Use of Label Switching With RSVP](#), [Multiprotocol Label Switching Architecture](#). There are no RFC's at present.

The deployment of MPLS by various vendors is still proprietary, and a large number of IP switch vendors announced to support MPLS in their products.

Contributions to section 10 by [Fujitsu Nexion](#).

Section 11 - What about 802.1p/Q?

The IEEE 802.1p and 802.1Q standards define how Ethernet switches can classify traffic in order to expedite delivery of time critical traffic. See the 802.1 working group page at <http://grouper.ieee.org/groups/802/1/index.html>.

By using 3 bits for precedence, IP has the same eight levels of prioritization--0 through 7--as 802.1p, 802.1Q and most of the other LAN topologies. By implementing a cross-mapping service, it becomes feasible to provide a one-to-one mapping between 802.1Q Ethernet frames and the IP precedence, letting you build a network that carries prioritization from one Ethernet LAN to another

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

across an IP-based WAN or ISP connection. Many router vendors say they will be implementing this functionality in the coming months as they roll out their 802.1Q-compliant products.

11.1 802.1p

IEEE 802.1p is a key enabler to QoS by enabling "Prioritized Ethernet" with up to 8 priorities in Ethernet and Token Ring networks. It enables Audio/Video traffic on switched Ethernet fabrics. The eight discrete priority levels range from the default of best effort, through excellent effort (a business-critical application, but tolerant of some delay), interactive multimedia (sensitive to delay or jitter), and reserved (highest priority). This is a Layer 2 (Data Link) priority setting, as opposed to the ToS and IP Precedence/CBQ bits, which are Layer 3 (Network Level) settings carried in the IPv4 header. Because it must be implemented in the hardware of network devices, existing switches and routers need to be replaced with ones supporting this technology. It is used in the Gigabit Ethernet standard to allow Ethernet to support Class of Service.

11.2 802.1Q

An IEEE standard for providing a virtual LAN capability within a campus network, used in conjunction with IEEE LAN protocols such as Ethernet and token ring. It describes Layer 2 VLAN markings (frame tagging). 802.1Q establishes a standard format for frame tagging, which will enable the creation of VLANs that use equipment from multiple vendors.

11.3 802.1D

802.1D bridges are based on a set of operational principles and an operational model that does not support the use of flow control mechanisms but can coexist with temporary congestion controls that protect limited input buffers.

Section 12 - What about QoS in Layer 2 and Layer 3 switching?

12.1 Layer 2

See section 11.1, 802.1P.

12.2 Layer 3

Layer-3 switches enable much of an IP router's functionality with the simplicity, efficiency, and lower-cost of a network switch. They use IP network layer information, as well as Layer 4 (TCP or UDP port numbers) or application attributes, to apply policy for traffic classification, filtering and forwarding. Like standard routers, Layer-3 switches do not mark/unmark traffic. One of the drawbacks to a campus network based entirely on layer-two switching is that reconfiguration, typically handled with the Spanning Tree protocol, tends to be slow. A network that uses layer-three switching in the backbone is able to more rapidly route around failures in cable, AC power, switch ports, or entire switches.

Section 13 - How does routing support QoS?

The IETF working group for QoS Routing is chartered to define a framework and techniques for Quality of Service (QoS) Routing in the Internet. QoS Routing allows the network to determine a path that supports the QoS needs of one or more flows in the network. The path chosen may not be the "traditional shortest path" that is typically computed based on current metrics and policies.

For example, QoS-enabled routers can support the ability of advanced applications to mark their packets for priority treatment. Marked packets are always sent first and are never dropped during congestion. A separate priority queue is maintained for marked packets and never oversold.

See the working group charter at <http://www.ietf.org/html.charters/qosr-charter.html> and their RFC [A Framework for QoS-based Routing in the Internet \(RFC 2386\)](#).

13.1 Diff-serv, Int-serv routing

TBS

13.2 Routing at ingress

TBS

13.3 Interior routing

TBS

13.4 Interdomain considerations

TBS

Section 14 - What are some implementation considerations?

14.1 Large Service providers and interdomain considerations

Reaching a common understanding about QoS mechanisms for inter-domain use, multi-vendor interoperability, and expected service behaviors in a network will be challenging.

Enabling QoS-which essentially allows one user to get a better service than another-requires policy enforcement, and that will require a policy management infrastructure. However, policy cannot be enforced unless the identities of network users can be

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

established to assign a level of trust. This implies a need for an authentication infrastructure. Since QoS provides added value there is a need for accounting and billing infrastructure. These three support services—Policy Management, Authentication, and Accounting/Billing—are essential to the success of QoS. All of them represent technical challenges that are being addressed, but more importantly, they represent significant new business opportunities that will provide further incentive to QoS deployment. Another very important scaling consideration that applies to all three of these services is managing peering arrangements between various Internet Service Providers (ISPs). Before enabling end-to-end QoS, bilateral agreements must be in place so ISPs sharing QoS responsibilities for common flows can share the necessary policy, authentication and accounting/billing information.

14.2 Within a customer network

TBS

14.3 Interoperability

The figure below shows how some of these QoS technologies can work together to provide "end-to-end QoS". [e2e-QoS]. Aside from the bandwidth broker—which is still a new concept at this point in time—this represents the model under development within the IETF community.

RSVP provisions resources for network traffic, whereas DiffServ simply marks and prioritizes traffic. RSVP is more complex and demanding than DiffServ in terms of router requirements, so can negatively impact backbone routers. This is why the "best common practice" says to limit RSVP's use on the backbone [AppStatement], and why DiffServ *can* exist there.

DiffServ is a perfect compliment to RSVP as the combination can enable end-to-end quality of service (QoS). End hosts may use RSVP requests with high granularity (e.g. bandwidth, jitter threshold, etc.). Border routers at backbone ingress points can then map those RSVP "reservations" to a class of service indicated by a DS-byte (or source host may set the DS-byte accordingly also). At the backbone egress point, the RSVP provisioning may be honored again, to the final destination. Ingress points essentially do traffic conditioning on a customer basis to assure that service level agreements (SLAs) are satisfied.

The architecture represented in the figure below —RSVP at the "edges" of the network, and DiffServ in the "core"—has momentum and support. Work within the IETF DiffServ work group is progressing quickly, although initial tests have shown mixed results.

14.4 QoS support in application software

TBS

14.5 Monitoring and measuring QoS

TBS

14.6 Managing QoS

TBS

14.7 QoS solutions for IP Multicast

TBS

14.8 Security and authentication

Work in the IETF public key infrastructure (PKIX) work group was recently finalized and since Internet commerce via SSL-enabled "web-shopping" has been using "certificates" for a number of years already, the technologies have had a chance to mature. There is still some work to be done, and things may yet change as wrinkles are ironed-out, but the certificate infrastructure to provide credentials for user/server/enterprise authentication and thereby enable trust on the Internet is good.

14.9 Billing, pricing and accounting

The work in this area has started in the IETF in the Authentication, Authorization and Accounting [AAA] working group. It would seem that a Bandwidth Broker third party could manage the billing and accounting for a number of ISPs, in addition to their SLAs. This concept is not unlike the current arrangements in the de-regulated telecommunications industry. Third parties broker on behalf of phone customers to find the best carriers and RBOCs (Regional Bell Operating Companies) for local and long-distance services at any point in time.

14.10 Testbeds (QBone)

The goal of the QBone is to provide an interdomain testbed for differentiated services (DiffServ), where the engineering, behavior, and policy consequences of new IP services can be explored. See the Internet2 Quality of Service page at <http://www.internet2.edu/qos>.

Section 15 - IP QoS over other network infrastructures

15.1 ATM

ATM and IP have to work together if the telephone network is to interoperate (and converge) with the Internet to enable seamless telephony services.

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.

ATM plays an important role in telephone network backbones, and its salient feature is "quality of service" (QoS) support. By allocating resources to a virtual circuit during connection setup that remain dedicated for the duration of the connection, ATM can satisfy the real-time (isochronous) delivery requirements of a two-way phone conversation.

The virtual circuit architecture of ATM is in stark contrast to the packet-switched design of IP, however. In addition to ATM's 53-byte "cell" size, and the fact that ATM is a data-link layer protocol as well as a network-layer like IP, these differences raise questions about compatibility. Fortunately, the work to ensure that IP can operate over ATM networks is done, and proven to work well. ATM's Available Bit Rate (ABR) service is actually intended to provide a service similar to IP's Best Effort. Current market indications are that ATM is unlikely to go to the desktop (i.e. to Internet hosts). This means that despite ATM's provisioning, at least part of the network will still only provide IP's "best effort" service. Hence, they are still susceptible to network congestion that can affect data throughput, and thereby adversely affect a telephony application. IP networks need a way to map to the QoS of ATM and extend it to the pure-IP portions of the Internet. To this end, work is underway to map ATM QoS to IP's RSVP.

See the IETF Integrated Services over Specific Link Layers ISSLL working group page at <http://www.ietf.org/html.charters/issll-charter.html> which has produced several RFC's addressing RSVP and ATM: [RSVP over ATM Implementation Guidelines \(RFC 2379\)](#), [Interoperation of Controlled-Load Service and Guaranteed Service with ATM \(RFC 2381\)](#), [A Framework for Integrated Services and RSVP over ATM \(RFC 2382\)](#) and [RSVP over ATM Implementation Requirements \(RFC 2380\)](#).

15.2 IP QoS over ATM using layer 3 switching in carrier networks

TBS

15.3 Frame relay

Ferguson's book [Quality of Service](#) has a chapter on QoS and Frame Relay, concluding with "You can construct an IP over Frame Relay network that adheres to QoS policies if you can modify the standard frame relay mode of operation."

Section 16 - Where can I find developer information?

16.1 Windows: Winsock2 and GQoS API

The Winsock 2 API provides RSVP support. The architecture supports a packet classifier to direct traffic to queues in the pack scheduler, which delivers queued packets to the network. See the QoS chapter in Quinn, Shute, [Windows Sockets Network Programming](#).

The Windows stack provides traffic control -queue management in stack, signaling - RSVP, packet marking - IP DiffServ and 802.1P, and admission control.

Section 17 - What's going on standards committees?

17.1 IETF

QoS Working Group Activities: DIFFSERV, RSVP, ISSLL (Integrated Services over Specific Link Layers), INTSERV (Integrated Services), MPLS

Policy Working Group Activities: POLICY, IPSP (IP Security Policy), RAP (RSVP Admission Policy)

Accounting Working Group Activities AAA (Authentication, Authorization & Accounting)

Telephony Working Group Activities: IPTEL (IP Telephony), MEGACO (Media Gateway Control) SIGTRAN (Signaling Transport), PIN (PSTN/Internet Notification), PINT (PSTN and Internet Internetworking), AVT Audio/Video Transport, MMUSIC (Multiparty Multimedia Session Control)

see www.ietf.org

17.2 ITU

TBS

Section 18 - Useful References

18.1 Books

Paul Ferguson, Geoff Huston, [Quality of Service : Delivering Qos on the Internet and in Corporate Networks](#)

Maufer, Tom, [Deploying Multicast in the Enterprise](#)

Quinn, Shute, [Windows Sockets Network Programming](#)

18.2 Mailing lists

TBS

18.3 Web sites

www.qosforum.com, www.ietf.org

18.4 IETF

www.ietf.org

To join the Quality of Service, contact [Nancy Moss](#) at (408)879-8080 or register with QoSF.