Quality of Service Over IP

Raj Jain

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

Overview

- Integrated services
- Resource Reservation Protocol: RSVP
- Differentiated Services
- QoS routing
- Multiprotocol Label Switching (MPLS) CoS

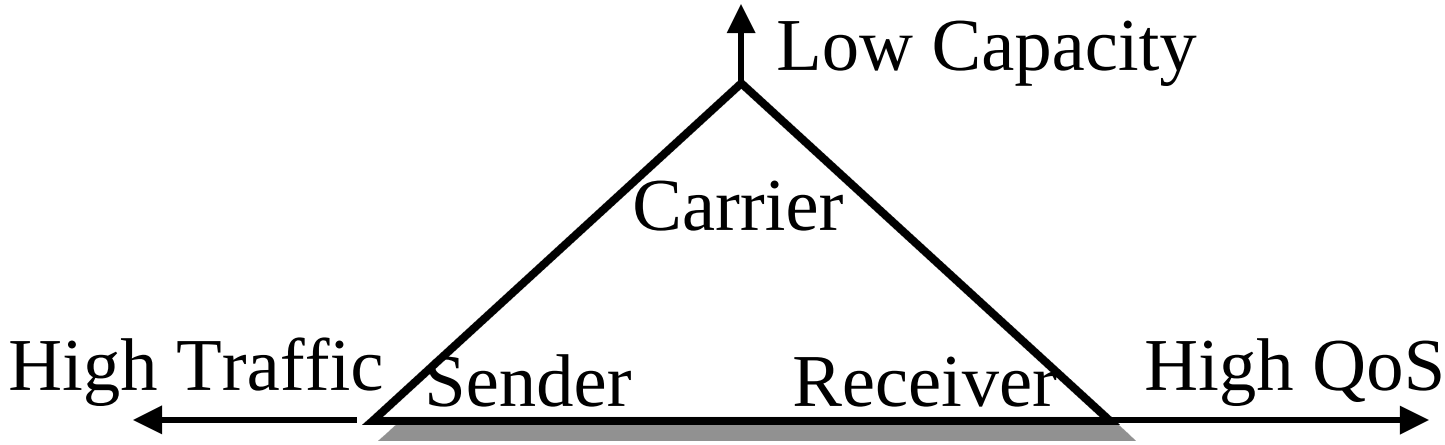
Multimedia

Application	Video Conferencing, Telephony, Fax
Transport	Timing Sync, payload id, error recovery
Network	QoS, Multicast, Signaling
Datalink	Access Control, Multicast, signaling
Physical	Multiple channels via SDM, FDM, TDM

IETF Groups

Application	Iptel, fax
Transport	Avt (RTP), mmusic (RTSP)
Network	Qosr, MPLS, IntServ, Issl, Diff-serv, RSVP
Datalink	IEEE 802.1p
Physical	Broadband Ethernet 10Broad36

QoS Triangle



- Senders want to send traffic any time with high load, high burstiness
- Receivers expect low delay and high throughput
- Since links are expensive, providers want to minimize the infrastructure
- If one of the three gives in ⇒ no problem

Components of QoS Architecture

1. Services with different QoS: Service definitions
2. Ways for users to communicate what they need: Signaling or admission control
3. Ways for providers to ensure that users are following their commitment: Policing/shaping
4. Ways for providers to find the routes: QoS based routing
5. QoS based forwarding: Buffer Allocation and Drop Policy, Queueing Discipline and Service Policy, Traffic Management of elastic traffic

ATM vs IP

QoS Component	ATM	IP
Services	CBR, VBR, ABR, UBR	Integrated Services
Signaling	UNI 4.0	RSVP
Policing/Shaping	Leaky bucket	Token bucket
Forwarding	Per-VC/per-class queueing	Differentiated Services
Elastic Traffic Mgmt	ABR	RED, Slow-start
Routing	PNNI	QoS routing

Integrated Services

- Best Effort Service
- Controlled-Load Service: Performance as good as in an unloaded datagram network. No quantitative assurances. (Min throughput)
- Guaranteed Service: rt-VBR
  - Firm bound on data throughput and delay.
  - Delay jitter or average delay not guaranteed or minimized.
  - Every element along the path must provide delay bound.
  - Is not always implementable, e.g., Shared Ethernet.

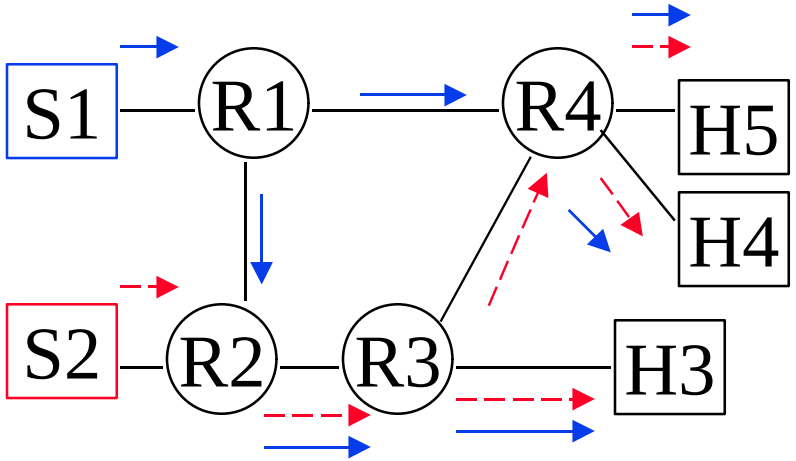
Flow Specification

- TSpec: Token bucket parameters
- RSpec (QoS): Allocated Rate (R) and delay slack (S)
  S = Extra acceptable delay over that obtainable with R
  Zero slack ⇒ Reserve exactly R.
- RSpec specified only for guaranteed rate service.
  Not for controlled load service.

RSVP

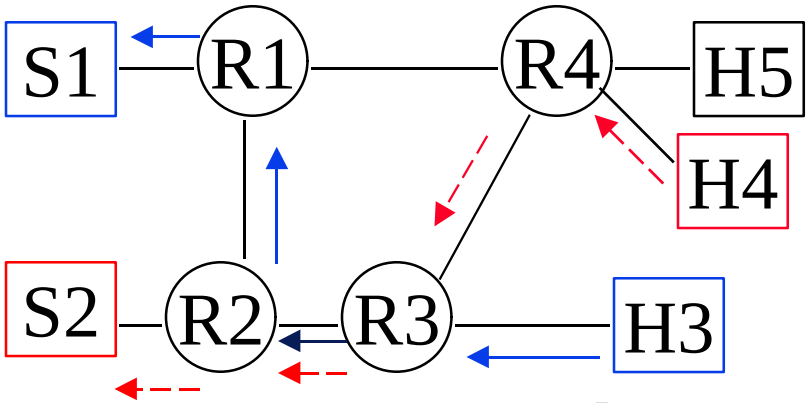
- Resource ReSerVation Protocol
- Internet signaling protocol
- Carries resource reservation requests through the network including traffic specs, QoS specs, network resource availability
- Sets up reservations at each hop
- RSVP does not find routes.
  Multicast routing protocols do.

Path Messages



- Sources send quasi-periodic PATH messages to multicast address
- Path message contain:
  - Sender Template: Data format, Src Address, Src Port
  - Sender TSpec: Traffic Characteristics. Not changed.
  - ADSpec: Network path resource/service availability Accumulated along the path.

Reservation Requests



- Receivers must join multicast address to receive path messages
- Receivers generate reservation (RESV) requests
- RESV messages contain resources to be reserved
- RESV messages are forwarded along the reverse path of PATH messages

Reservation (Cont)

- Requests are checked for resource availability (admission control) and administrative permissions (policy control)
- Two or more RESV messages for the same source over the same link are merged.
- Routers maintain a soft state. The receivers have to refresh periodically.
- Heterogeneous Receivers: Sources divide traffic into several flows. Each flow is a separate RSVP flow. Receivers join one or more flows. Each RSVP flow is homogeneous.

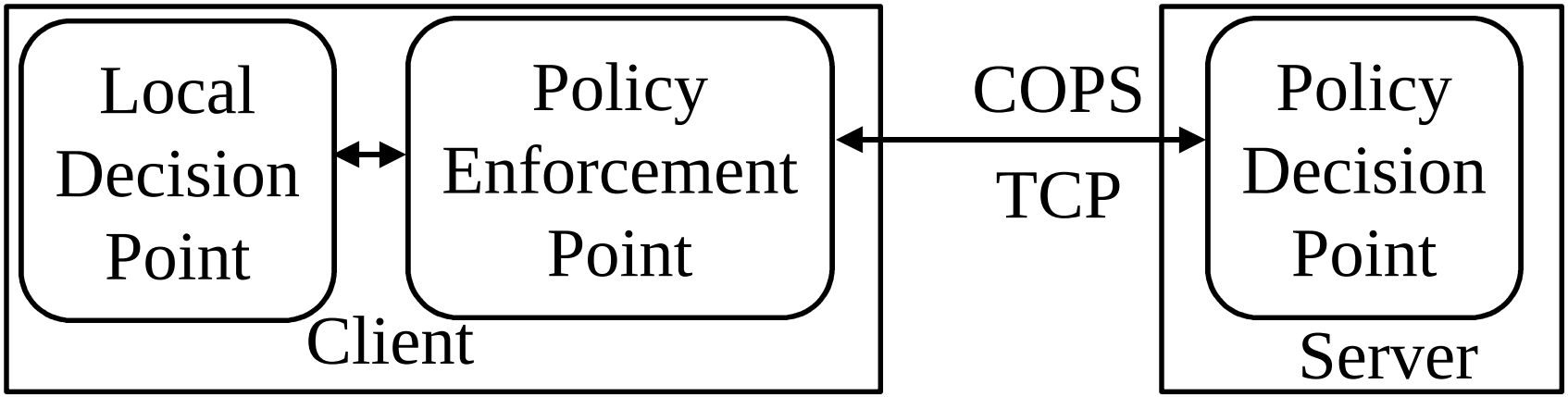
Problems with RSVP and Integrated Services

- Complexity in routers: packet classification, scheduling
- Per-Flow State: O(n) $\Rightarrow$ Not scalable.
  Number of flows in the backbone may be large.
  $\Rightarrow$ Suitable for small private networks
- Need a concept of "Virtual Paths" or aggregated flow groups for the backbone
- Need policy controls: Who can make reservations?
  Support for accounting and security.
  $\Rightarrow$ RSVP admission policy (rap) working group.

Problems (Cont)

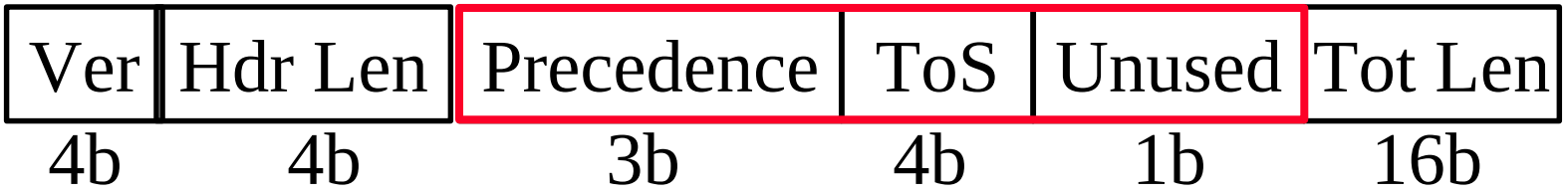
- Receiver Based:
  Need sender control/notifications in some cases.
  Which receiver pays for shared part of the tree?
- Soft State: Need route/path pinning (stability).
  Limit number of changes during a session.
- Throughput and delay guarantees require support of lower layers. Shared Ethernet $\Rightarrow$ IP can't do GS or CLS. Need switched full-duplex LANs.
- Can't easily do RSVP on ATM either
- Most of these arguments also apply to integrated services.

COPS Protocol



- Common Open Policy Service Protocol
- When the routers (clients) receive a RSVP message, they send the request the server and obtain authorization
- Will work with other (non-RSVP) signaling
- Routers can make local decisions but should keep servers informed
- Servers can send unsolicited responses for changes later

IP ToS Field



- IPv4: 3-bit precedence + 4-bit ToS
- RFC791: ToS determines packet treatment and monitory considerations
- RFC1349: $bit_1 \Rightarrow$ min delay, $bit_2 \Rightarrow$ max throughput, $bit_3 \Rightarrow$ max reliability, $bit_4 \Rightarrow$ min cost
- OSPF and integrated IS-IS can compute paths for each ToS

Differentiated Services Working Group

- August 97: BOF started
- Feb 98: Working group formed
- Dec 98: Final document
- Email: majordomo@baynetworks.com in body: subscribe diff-serv
- Archive: http://www-nrg.ee.lbl.gov/diff-serv-arch/
- Charter: define ds byte (IPv4 ToS or IPv6 traffic class octets)

Diff-Serv Terminology

- Service: Offered by the protocol layer
  - Application: Mail, FTP, WWW, Video,...
  - Transport: Delivery, Express Delivery,... Best effort, controlled load, guaranteed service
- Per-Hop Behavior (PHB): Mechanisms - Drop threshold, Queue assignment, Service priority, Service Rate
- Flow: Packets with specific header fields, Destination Address, Source Address, Port, Flow Label
- Aggregates: Stream of packets with the same DS byte pattern

Initial proposals

- Assured service (Jacobson): traffic profile (VBR or CLS like), in-profile and out-profile
- Premium Service (Clark): Peak rate (CBR or GS like), Virtual leased line
- Two-bit Service: A-bit (CLP) and P-bit (Priority)
- 2 Priority bits, 1 drop bit
- Bits for delay class: 2 bits $\Rightarrow$ 4 classes
  Bits for Drop preference: 3 bits $\Rightarrow$ 8 classes

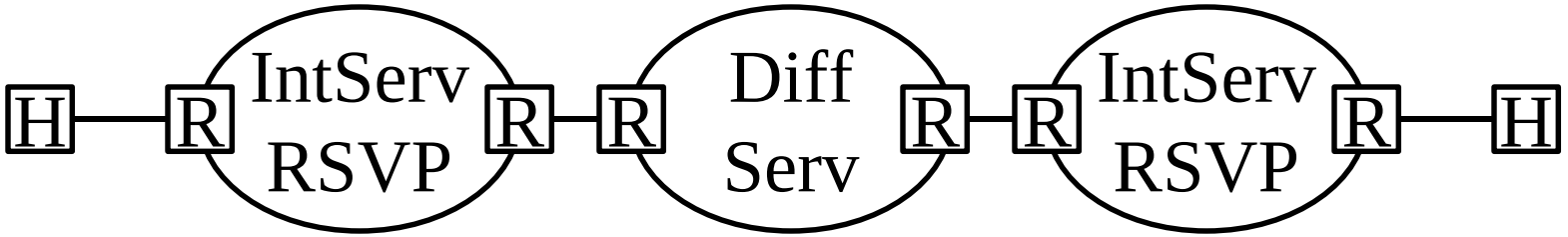
Latest PHB Allocation

- ppp i 00
- ppp = Precedence (Higher is generally better)
- i = in/out bit ⇒ In profile/out Profile
  ⇒ Drop preference. Allows in/out pkts in same Queue
- Compatible with current usage
- Precedence is used as an index to select a queue, or VC, ...
- In IEEE-802 switches, only 1, 2, or 3 msbs used
- Unrecognized code points ⇒ Default forwarding

PHB Allocation (Cont)

- Plan: 32 code points standard,
  16 Experimental/local use, 16 reserved
  xxxxx0 Standard
  xxxx11 Experimental/Local Use
  xxxx01 Reserved for future

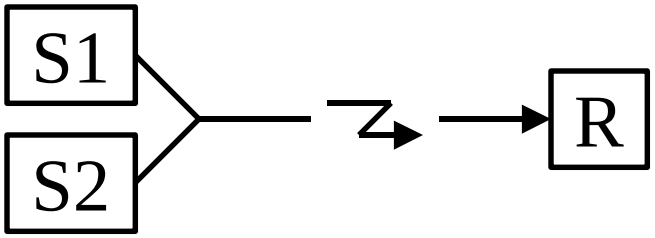
End-to-end QoS



- Hosts may mark DS byte or use RSVP signaling or both or none.
- Why hosts? 1. Encryption, 2. Hosts know the importance of info even if the header fields are same
- Routers may mark DS byte if necessary.
- Routers at the intserv diff-serv boundary accept/reject RSVP requests based on current load

- Service between intserv and diff-serv regions can be statically or dynamically provisioned
- Current integrated services (CLS, GS) may or may not be practical
- DS byte may be modified at network boundary

Issues



- Standard code points (behaviors)
- Receiver control over incoming low-speed link
- Signaling: Should users signal or network managers set resource allocations
- Dynamic or Static management controls?
- Billing: Bit for receiver billing. If receiver billing, the receiver should be able to deny/drop packets received.
- Congestion Check Bit: If set, network indicates highest priority for which packets are being dropped in the ToS byte.

QoS Extensions to OSPF

- Open shortest path first
- Separate metric can be specified for each ToS supported
- OSPF options field has a T-bit
  T-bit = 1 ⇒ Router can compute routes for each ToS
- Work to extend OSPF is currently underway
- QoS ⇒ Frequent updates
  ⇒ Instability: Underloaded links become overloaded
  Also, complexity
- **Ref**: Z. Zhang, et al, "QoS Extensions to OSPF," Sep. 97, http://www.internic.net/internet-drafts/draft-zhang-qos-ospf-01.txt

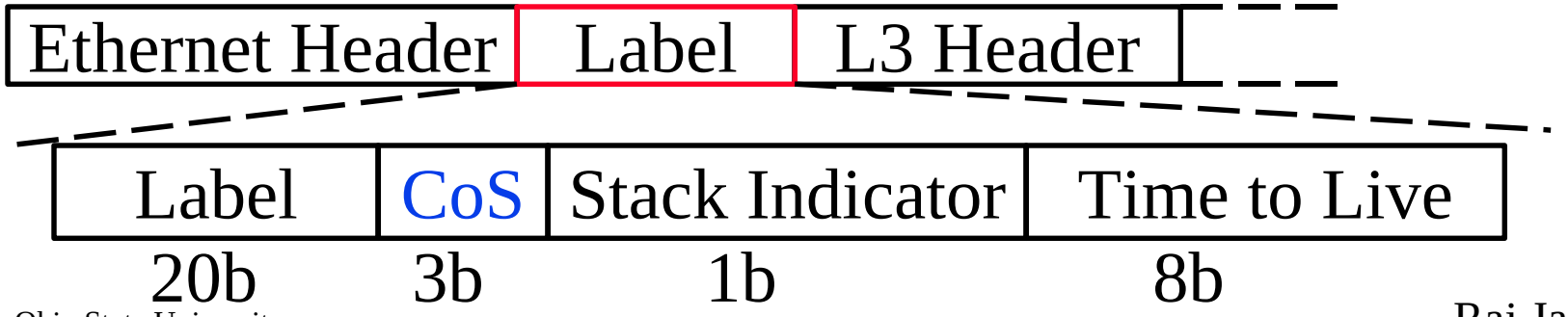
Inter-Domain QoS Routing

- Domains want to limit the frequency and amount of information exchanged $\Rightarrow$ Stability
- QoS based routing may cause frequent changes and instability
- QoS extensions to Border Gateway Protocol (BGP) proposed but may or may not happen
- Need hierarchical aggregation for scalability
  Crank-back

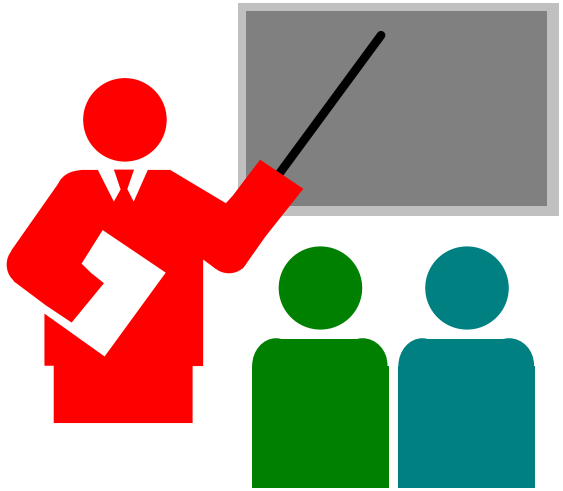
MPLS

Prefix	Label	Out-Port	Out-Label
164.107.0.0/16	1	2	3
164.107.0.0/24	2	3	4
…	…	…	…

- Multiprotocol Label Switching
- Current: Longest prefix match on the dest address
- With Labels: Search can be replaced by indexing
- MPLS labels contain 3-bit CoS



Summary



- Internet protocols suite is being extended to allow QoS
- Integrated Services: GS = rtVBR, CLS = nrt-VBR
- Signaling protocol: RSVP
- Differentiated Services will use the DS byte
- QoS Routing: QOSPF
- Multiprotocol Label Switching has 3-bit CoS

References

- For a detailed list of references see: http://www.cis.ohio-state.edu/~jain/refs/ipqs_ref.htm
- P. Ferguson and G. Huston, "Quality of Service: Delivering QoS on the Internet and in Corporate Networks," Wiley 1998.
- RFC 2212, "Specification of Guaranteed Quality of Service", 9/97
- RFC 2211 "Specification of the Controlled-Load Network Element Service", 9/97

References (Cont)

- RFC 2205, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", 11/97
- "The COPS (Common Open Policy Service) Protocol", 03/16/1998, draft-ietf-rap-cops-01.txt
- "A Framework for Multiprotocol Label Switching", 11/26/1997, draft-ietf-mpls-framework-02.txt
- "MPLS Label Stack Encoding", 02/18/1998, draft-ietf-mpls-label-encaps-01.txt
- "A Framework for QoS-based Routing in the Internet", 04/09/1998, draft-ietf-qosr-framework-04.txt

References (Cont)

- Diff-Serv Mail Thread index, http://www-nrg.ee.lbl.gov/diff-serv-arch/
- "Definition of Differentiated Services Field (DS Byte) in the IPv4 and IPv6 Headers," 5/1998, draft-ietf-diffserv-header-00.txt
- "Differentiated Services Operational Model and Definitions", 02/11/1998, draft-nichols-dsopdef-00.txt
- "IP Precedence in Differentiated Services Using the Assured Service", 04/10/1998, draft-ietf-diffserv-precedence-00.txt

References (Cont)

- "A Framework for End-to-End QoS Combining RSVP/IntServ and Differentiated Services", 03/16/1998, draft-bernet-intdiff-00.txt

IETF Working Groups

- Internet Fax (fax)
- IP Telephony (iptel)
- Audio/Video Transport (avt)
- MBONE deployment working group (mboned)
- Multiparty Multimedia Session Control (mmusic)
- Multicast Extensions to OSPF (mospf)
- Inter-Domain Multicast Routing (idmr)
- Large Scale Multicast Applications (lsma)
- Integrated Services (intserv)
- Integrated Services over Specific Link Layers (issll)
- Resource Reservation Setup Protocol (rsvp)
- QoS-based Routing (qosr)
- Differentiated services (diff-serv)

List of Acronyms

ABR	Available Bit Rate
ATM	Asynchronous Transfer Mode
BA	Behavior Aggregate
BGP	Border Gateway Protocol
BOF	Birds of a Feather
CBR	Constant Bit Rate
CDV	Cell Delay Variation
CFI	Canonical Format Indicator
CLP	Cell Loss Priority
CLS	Controlled Load Service
COPS	Common Open Policy Service Protocol

Acronyms (Cont)

CoS	Class of Service
DA	Destination Address
DQDB	Distributed Queue Dual Bus
DSBM	Designated Subnet Bandwidth Manager
DVMRP	Distance Vector Routing Multicast Protocol
FCS	Frame Check Sequence
FDDI	Fiber Distributed Data Interface
FIFO	First in First out
FTP	File Transfer Protocol
GS	Guaranteed Service
ICMP	Internet Control Message Protocol

Acronyms (Cont)

IEEE	Institution of Electrical and Electronic Engineers
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IP	Internet Protocol
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
IS	Internal System
IntServ	Integrated Services
LANs	Local Area Networks
LLC	Logical Link Control
LU	Local Use

Acronyms (Cont)

MAC	Media Access Control
MBONE	Multicast Backbone
MBS	Maximum Burst Size
MF	Multi-field
MPLS	Multiprotocol Label Switching
MTU	Maximum Transmission Unit
NHRP	Next Hop Resolution Protocol
OOPS	Open Outsourcing Policy Service
OSPF	Open Shortest Path First
PASTE	Provider Architecture for Differentiated Services and Traffic Engineering

Acronyms (Cont)

PCR	Peak Cell Rate
PHB	Per-Hop Behavior
PIM	Protocol Independent Multicast
PT	Protocol Type
QOSPF	QoS-OSPF
QoS	Quality of Service
RED	Random Early Discard
ResV	Reservation Request
RFC	Request for Comment
RIF	Routing Information Field
RSVP	Resource Reservation Protocol

Acronyms (Cont)

RSpec	QoS Specification
RTP	Real-time Transport Protocol
SBM	Subnet Bandwidth Manager
SONET	Synchronous Optical Network
TCP	Transmission Control Protocol
TPID	Tag Protocol ID
TR	Token Ring
TSpec	Traffic Specification
ToS	Type of Service
UBR	Unspecified Bit Rate
UDP	User Datagram Protocol

Acronyms (Cont)

UNI	User-Network Interface
VBR	Variable Bit Rate
VC	Virtual Circuit
VLAN	Virtual Local Area Network
WAN	Wide Area Network
WFQ	Weighted Fair Queueing