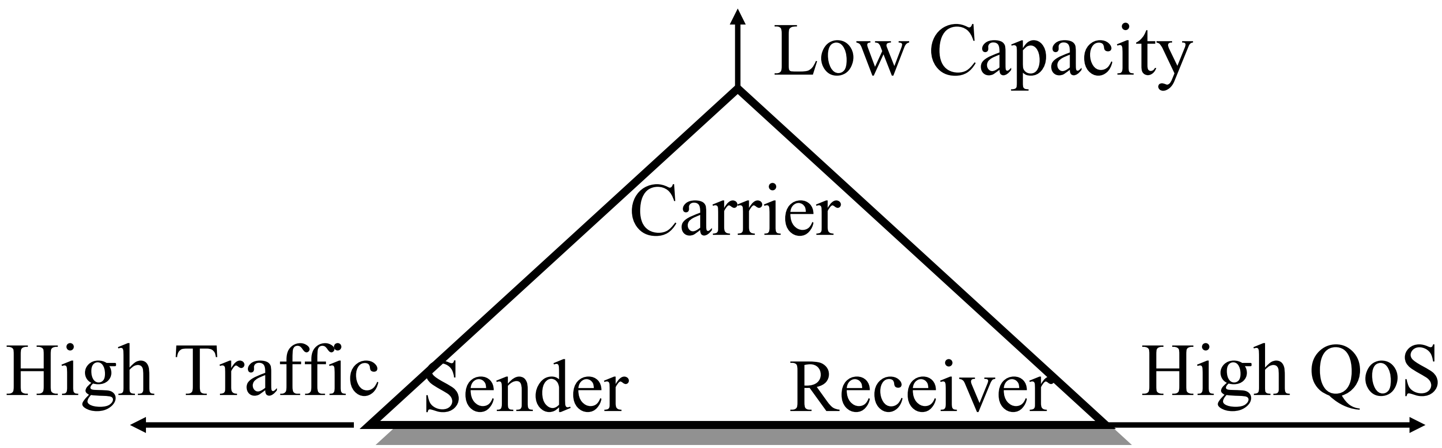
Quality of Service In Data Networks: Problems, Solutions, and Issues

Raj Jain, Panel Chair

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

- ATM QoS and Issues
- Integrated services/RSVP and Issues
- Differentiated Services and Issues
- QoS using MPLS
- End-to-end QoS
- This is an update to the May'98 talk
  http://www.cis.ohio-state.edu/~jain/talks/ipqos.htm

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

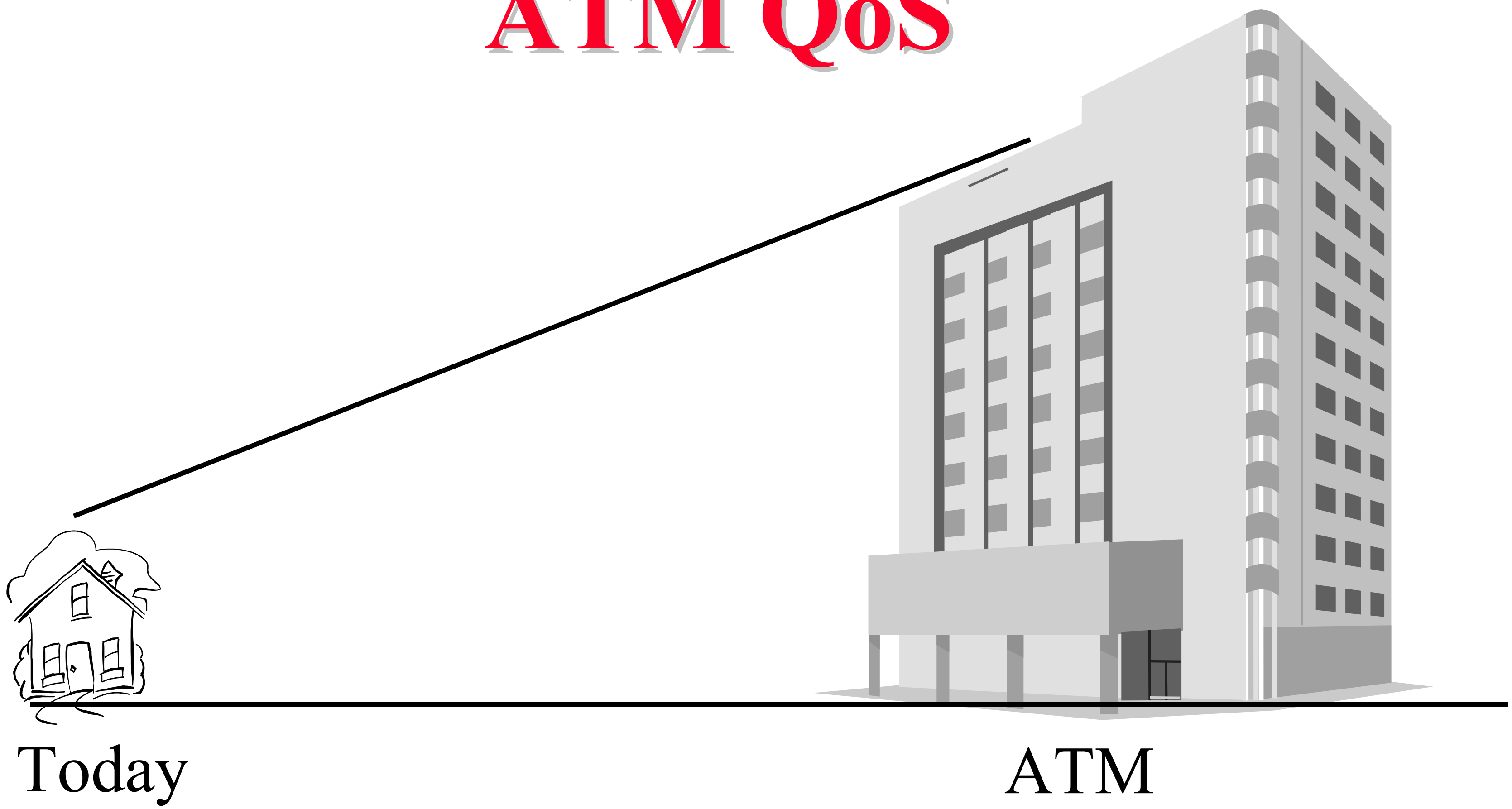
What is QoS?

- Predictable Quality: Throughput, Delay, Loss, Delay jitter, Error rate
- Opposite of best effort = Random quality
- Mechanisms:
  - Capacity Planning
  - Classification, Queueing, Scheduling, buffer management
  - QoS based path determination, Route pinning
  - Shaping, policing, admission control
  - Signaling

ATM Service Categories

- **CBR**: Throughput, delay, delay variation
- **rt-VBR**: Throughput, delay, delay variation
- **nrt-VBR**: Throughput
- **UBR**: No Guarantees
- **GFR**: Minimum Throughput
- **ABR**: Minimum Throughput. Very low loss. Feedback.
- ATM also has QoS-based routing (PNNI)

ATM QoS



Too much too soon

ATM QoS: Issues

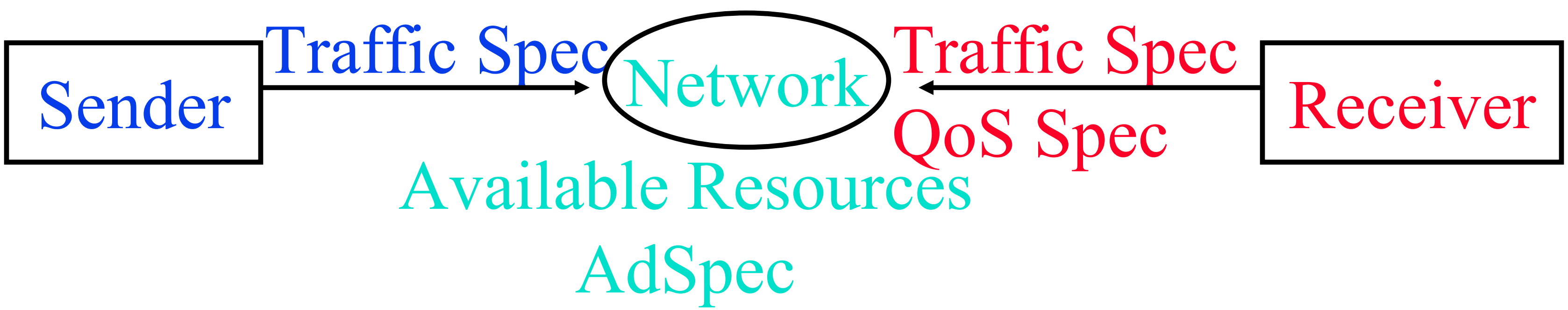
- Can't easily aggregate QoS: VP = Σ VCs
- Can't easily specifiy QoS: What is the CDV required for a movie?
- Signaling too complex $\Rightarrow$ Need Lightweight Signaling
- Need Heterogeneous Point-to-Multipoint: Variegated VCs
- Need QoS Renegotiation
- Need Group Address
- Need priority or weight among VCs to map DiffServ and 802.1D

Integrated Services

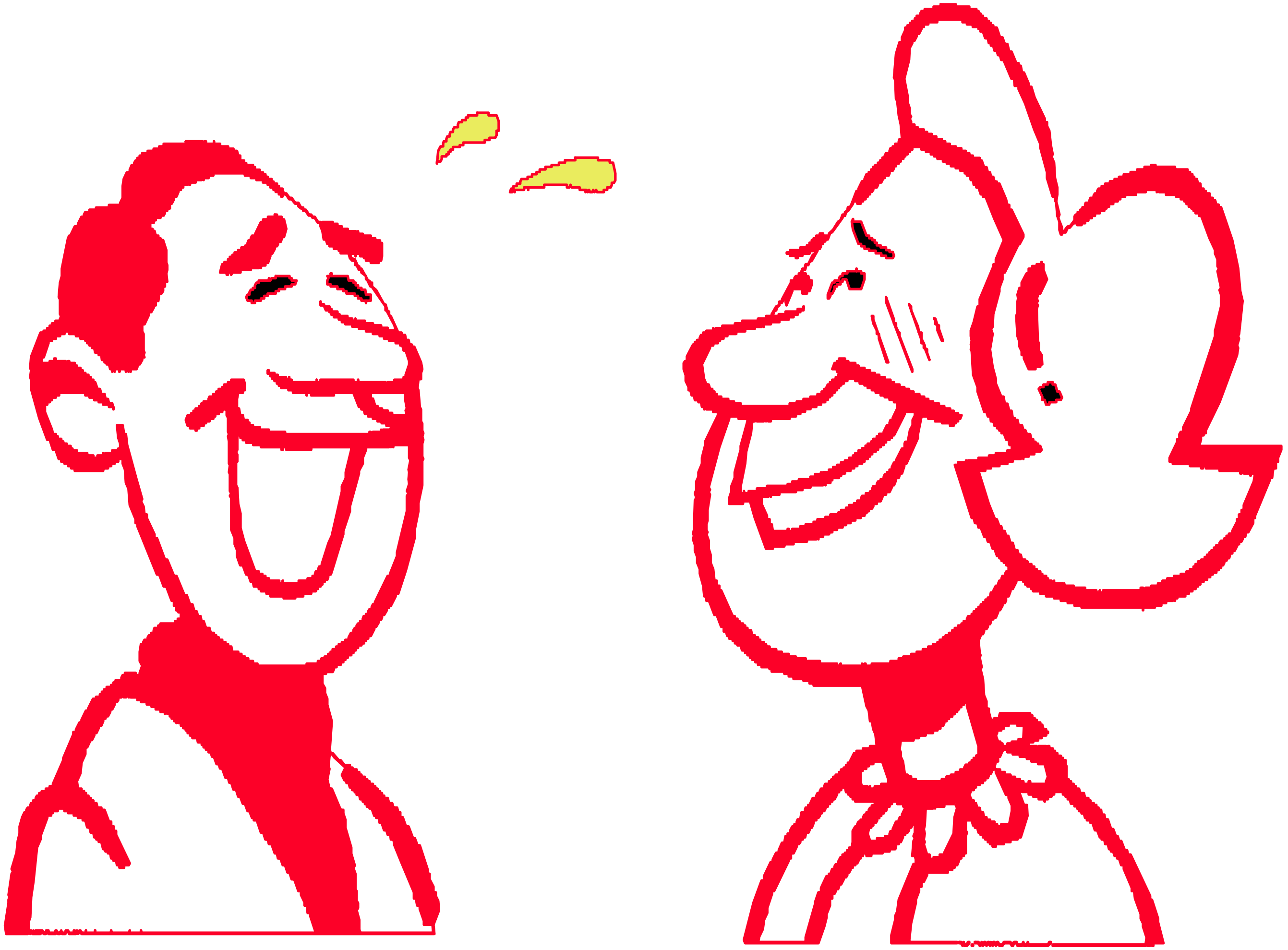
- Best Effort Service: Like UBR.
- Controlled-Load Service: Performance as good as in an unloaded datagram network. No quantitative assurances. Like nrt-VBR or UBR w MCR
- Guaranteed Service: rt-VBR
  - Firm bound on data throughput and delay.
  - Delay jitter or average delay not guaranteed or minimized.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

RSVP

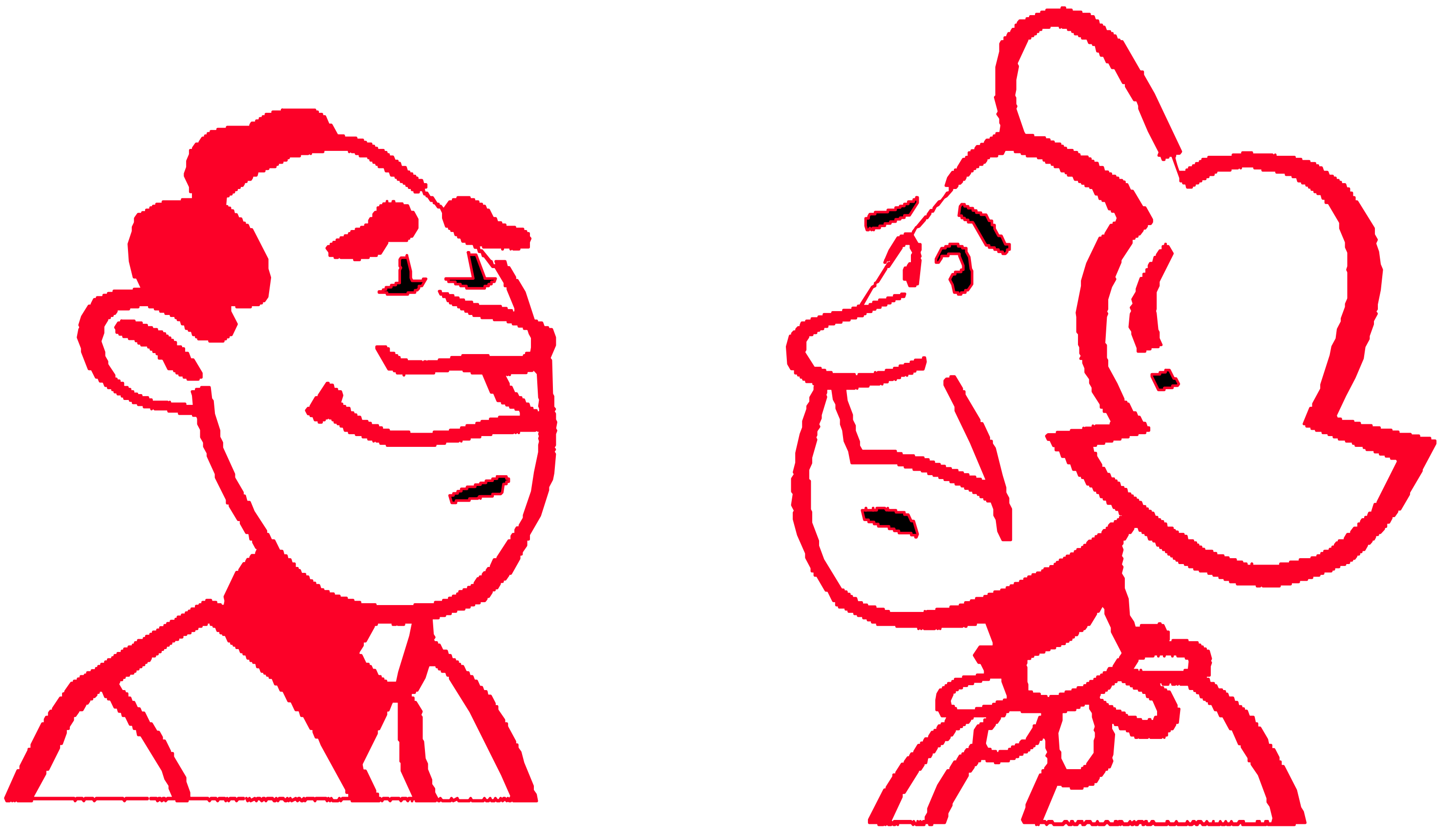
- Resource ReSerVation Protocol
- Internet signaling protocol
- Carries resource reservation requests through the network including traffic specs, QoS specs, network resource availability
- Sets up reservations at each hop



Before



After



Problems with RSVP and Integrated Services

- Complexity in routers: packet classification, scheduling
- Scalable in number of receivers per flow but Per-Flow State: O(n) ⇒ Not scalable with # of flows. Number of flows in the backbone may be large. ⇒ 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. ⇒ 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.
- RSVP does not have negotiation and backtracking
- Throughput and delay guarantees require support of lower layers. Shared Ethernet ⇒ 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.

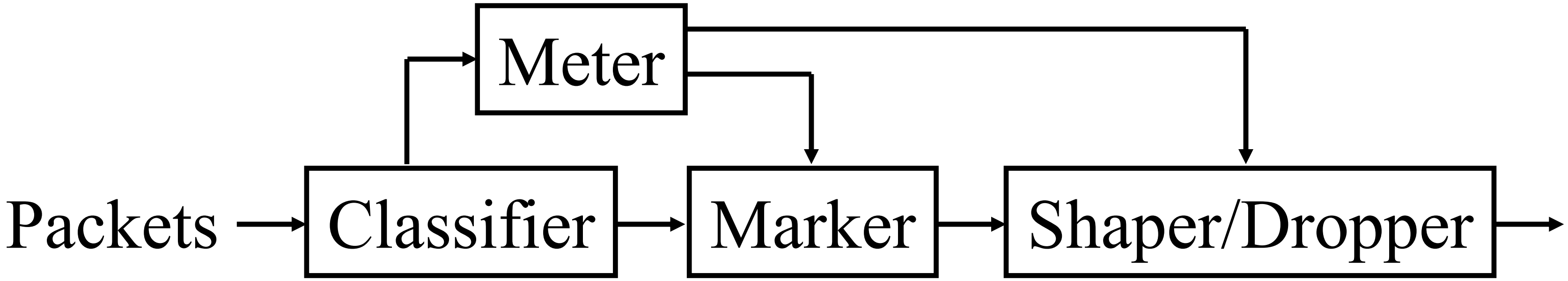
Differentiated Services

Ver	Hdr Len	Precedence	ToS	Unused	Tot Len
4b	4b	3b	4b	1b	16b

- IPv4: 3-bit precedence + 4-bit ToS
- OSPF and integrated IS-IS can compute paths for each ToS
- Many vendors use IP precedence bits but the service varies ⇒ Need a standard ⇒ Differentiated Services
- DS working group formed February 1998
- Charter: Define ds byte (IPv4 ToS field)
- Mail Archive: http://www-nrg.ee.lbl.gov/diff-serv-arch/

DiffServ Concepts

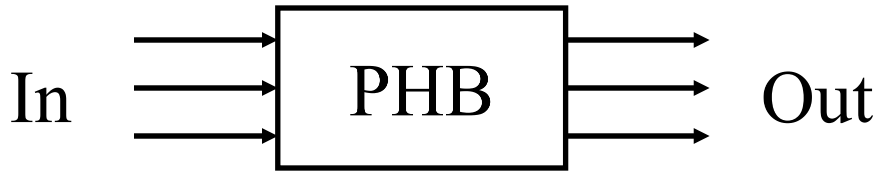
- Micro-flow = A single application-to-application flow
- Traffic Conditioners: Meters (token bucket), Markers (tag), Shapers (delay), Droppers (drop)
- Behavior Aggregate (BA) Classifier: Based on DS byte only
- Multi-field (MF) Classifiers: Based on IP addresses, ports, DS-byte, etc..



Diff-Serv Concepts (Cont)

- Service: Offered by the protocol layer
  - Application: Mail, FTP, WWW, Video,...
  - Transport: Delivery, Express Delivery,...
    Best effort, controlled load, guaranteed service
  - DS group will not develop services
    They will standardize "Per-Hop Behaviors"

Per-hop Behaviors

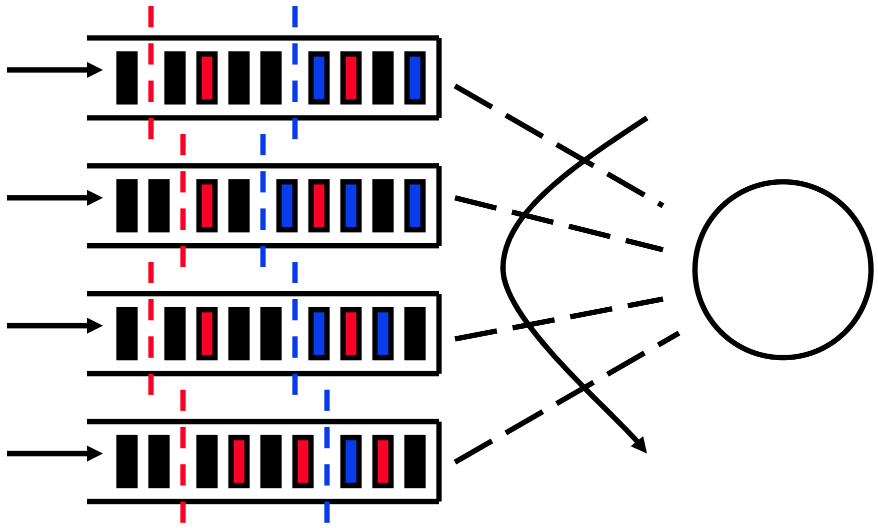


- Externally Observable Forwarding Behavior
- x% of link bandwidth
- Minimum x% and fair share of excess bandwidth
- Priority relative to other PHBs
- PHB Groups: Related PHBs. PHBs in the group share common constraints, e.g., loss priority, relative delay

Expedited Forwarding

- Also known as "Premium Service"
- Virtual leased line
- Similar to CBR
- Guaranteed minimum service rate
- Policed: Arrival rate < Minimum Service Rate
- Not affected by other data PHBs
  $\Rightarrow$ Highest data priority (if priority queueing)
- Code point: 101 110

Assured Forwarding



- PHB Group
- Four Classes: No particular ordering
- Three drop preference per class

Assured Forwarding (Cont)

- DS nodes SHOULD implement all 4 classes and MUST accept all 3 drop preferences. Can implement 2 drop preferences.
- Similar to nrt-VBR/ABR/GFR
- Code Points:

Drop Prec.	Class 1	Class 2	Class 3	Class 4
Low	010 000	011 000	100 000	101 000
Medium	010 010	011 010	100 010	101 010
High	010 100	011 100	100 100	101 100

- Avoids 11x000 (used for network control)

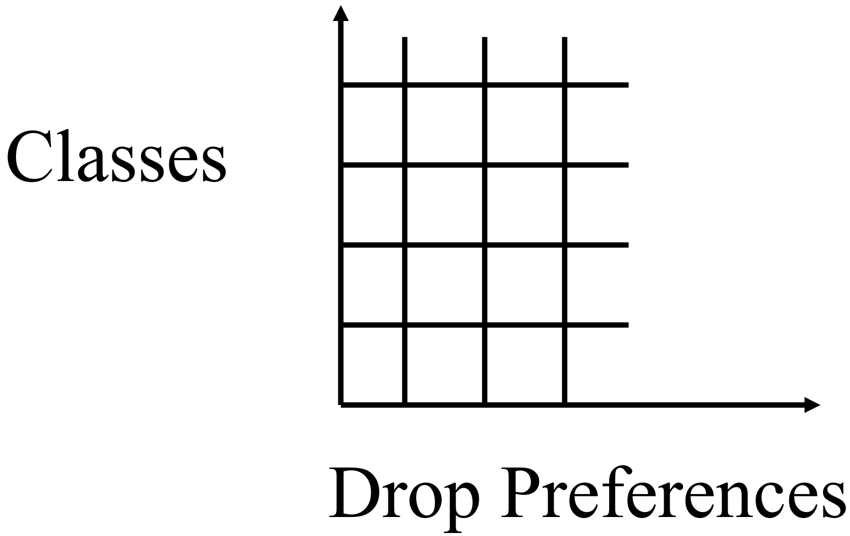
AF Simulation Results

1. W/O DPs, TCP is punished for good behaviour
2. Fairness is also poor.
3. Three DPs give the same perf for TCP as two DPs

Reason: TCP does not distinguish between loss of packets of different drop precedences

Reference: M. Goyal, et al, "Effect of Number of Drop Precedences in Assured Forwarding," IETF draft-goyal-dpstdy-diffserv-00.txt, March 1999, http://www.cis.ohio-state.edu/~jain/ietf/dpstdy.htm

On Drop Preferences



- We have two dimensions of control
  - Classes = Queues
  - Drop Preferences = Right to enter the queue
- Classes ⇒ Directly controls bandwidth allocation

Drop Preferences (Cont)

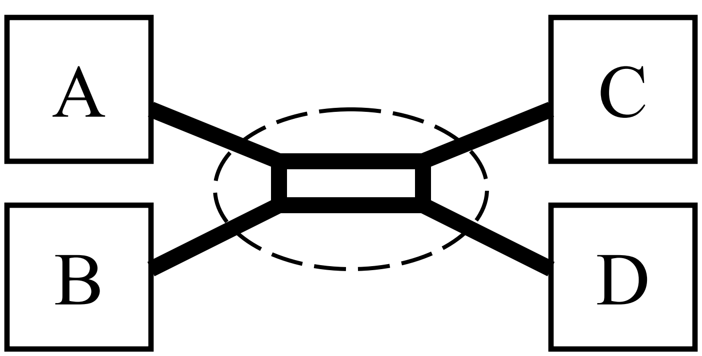
- DPs $\Rightarrow$ Controls buffer allocation
  $\Rightarrow$ Indirectly affects bandwidth allocation
  - Depends upon the arrival pattern
    $\Rightarrow$ Random $\Rightarrow$ Not Reliable
- Given a limited number of PHB's, it is better to have more classes than more DPs

Problems with DiffServ

- per-hop ⇒ Need at every hop
  One non-DiffServ hop can spoil all QoS
- End-to-end ≠ Σ per-Hop
  Designing end-to-end services with weighted guarantees at individual hops is difficult.
  Only EF will work.
- Designed for static Service Level Agreements (SLAs)
  Both the network topology and traffic are highly dynamic.
- Multicast ⇒ Difficult to provision
  Dynamic multicast membership ⇒ Dynamic SLAs?

DiffServ Problems (Cont)

- DiffServ is unidirectional ⇒ No receiver control
- Modified DS field ⇒ Theft and Denial of service. Ingress node should ensure.
- How to ensure resource availability inside the network?
- QoS is for the aggregate not per-destination. Multi-campus enterprises need inter-campus QoS.



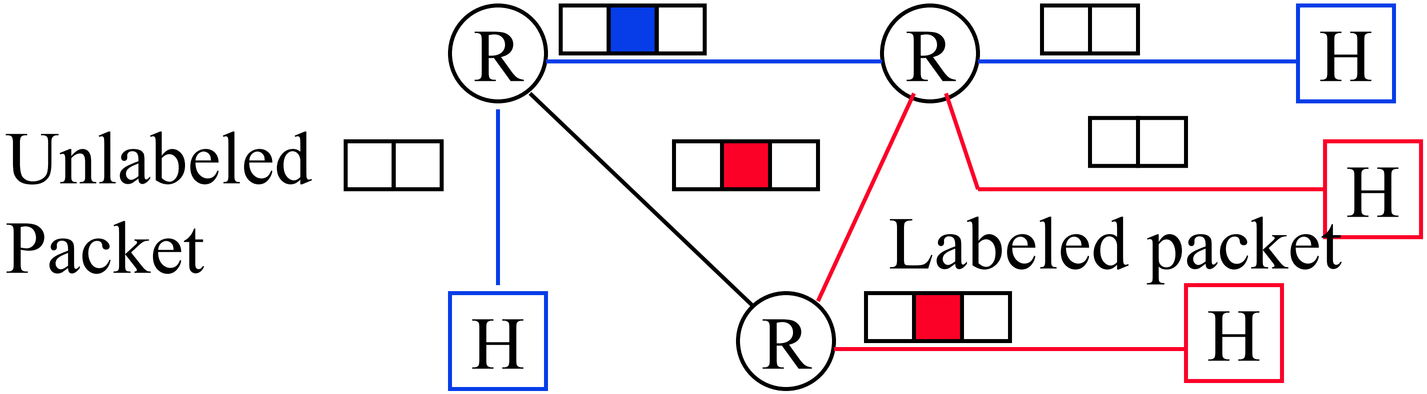
DiffServ Problems (Cont)

- QoS is for the aggregate not micro-flows. Not intended/useful for end users. Only ISPs.
  - Large number of short flows are better handled by aggregates.
  - Long flows (voice and video sessions) need per-flow guarantees.
  - High-bandwidth flows (1 Mbps video) need per-flow guarantees.
- All IETF approaches are open loop control ⇒ Drop
  Closed loop control ⇒ Wait at source
  Data prefers waiting ⇒ Feedback

DiffServ Problems (Cont)

- Guarantees $\Rightarrow$ Stability of paths
  $\Rightarrow$ Connections (hard or soft)
  Need route pinning or connections.

Multiprotocol Label Switching



- Entry "label switch router (LSR)" attaches a label to the packet based on the route
- Other LSRs switch packets based on labels. Do not need to look inside ⇒ Fast.
- Labels have local significance ⇒ Different label at each hop (similar to VC #)
- Exit LSR strips off the label

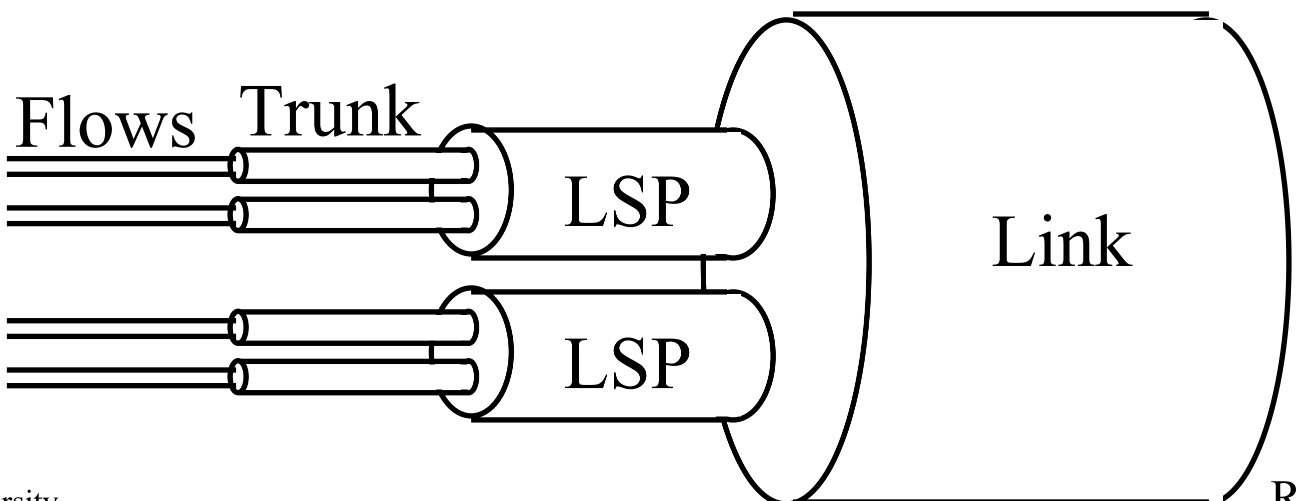
Traffic Engineering Using MPLS

- Traffic Engineering = Performance Optimization
  = Efficient resource allocation, Path splitting
  $\Rightarrow$ Maximum throughput, Min delay, min loss
  $\Rightarrow$ Quality of service
- In MPLS networks: "Traffic Trunks" = SVCs
  Traffic trunks are routable entities like VCs
- Multiple trunks can be used in parallel to the same egress.
- Each traffic trunk can have a set of associated characteristics, e.g., priority, preemption, policing, overbooking

Flows, Trunks, LSPs, and Links

- Label Switched Path (LSP):
  All packets with the same label
- Trunk: Same Label+Exp
- Flow: Same MPLS+IP+TCP headers

DL	Label	Exp	SI	TTL	IP	TCP	



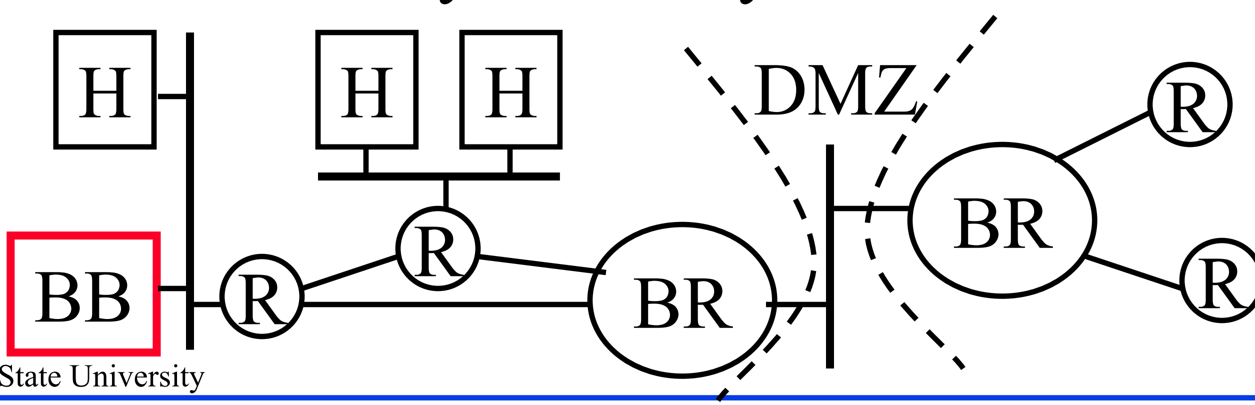
MPLS Simulation Results

- Total network throughput improves significantly with proper traffic engineering
- Congestion-unresponsive flows affect congestion-responsive flows
  - Separate trunks for different types of flows
- Trunks should be end-to-end
  - Trunk + No Trunk = No Trunk

Reference: P. Bhaniramka, et al, "*QoS using Traffic Engineering over MPLS: An Analysis*," IETF draft-bhani-mpls-te-anal-00.txt, March 1999, http://www.cis.ohio-state.edu/~jain/teanal.htm

Bandwidth Broker

- Repository of policy database. Includes authentication
- Users request bandwidth from BB
- BB sends authorizations to leaf/border routers
  Tells what to mark.
- Ideally, need to account for bandwidth usage along the path
- BB allocates only boundary or bottleneck



IEEE 802.1D Model

Dest Addr	Src Addr	Tag Prot ID	Pri	CFI	VLAN ID
		← 802.1Q header →			

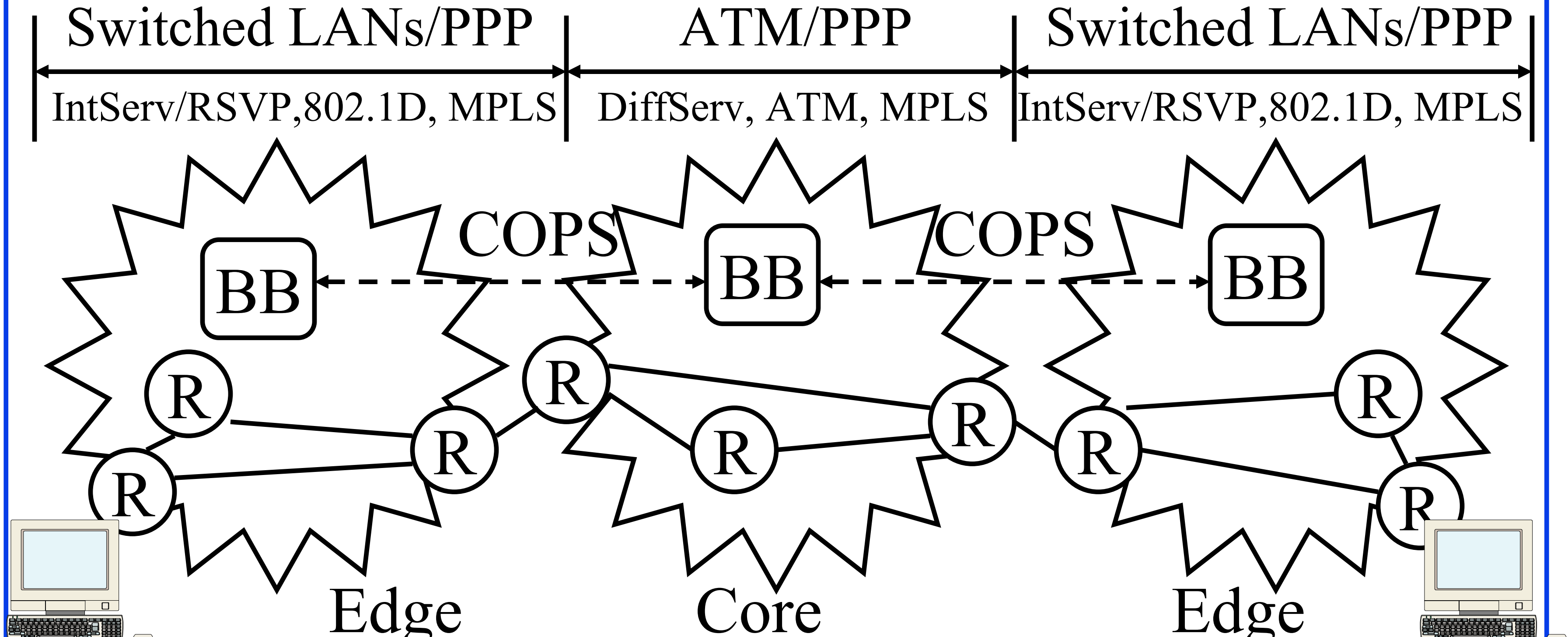
	Prot Type	Payload	FCS

CFI = Canonical Format Indicator (Source Routing)

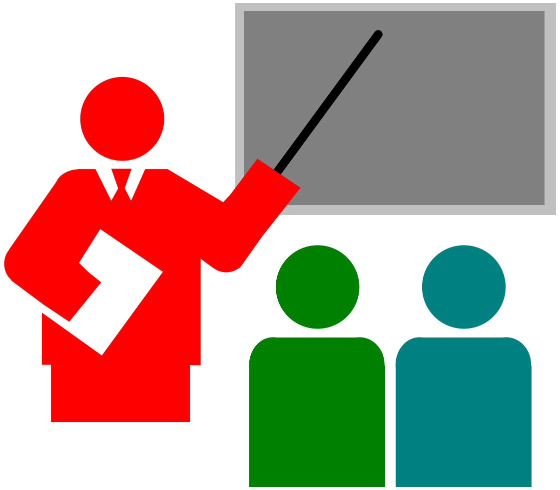
- **Up to eight priorities:** Strict.
  - 1 Background
  - 2 Spare
  - 0 Best Effort
  - 3 Excellent Effort
  - 4 Control load
  - 5 Video (Less than 100 ms latency and jitter)
  - 6 Voice (Less than 10 ms latency and jitter)
  - 7 Network Control

End-to-end View

- ATM/PPP backbone, Switched LANs/PPP in Stub
- IntServ/RSVP, 802.1D, MPLS in Stub networks
- DiffServ, ATM, MPLS in the core



Summary



- ATM: CBR, VBR, ABR, UBR, GFR
- Integrated Services: GS = rtVBR, CLS = nrt-VBR
- Signaling protocol: RSVP
- Differentiated Services will use the DS byte
- MPLS allows traffic engineering and is most promising
- 802.1D allows priority

References

- For a detailed list of references see: refs/ipqs_ref.htm
- Additional papers and presentations on QoS are at: http://www.cse.ohio-state.edu/~jain/

Thank You!