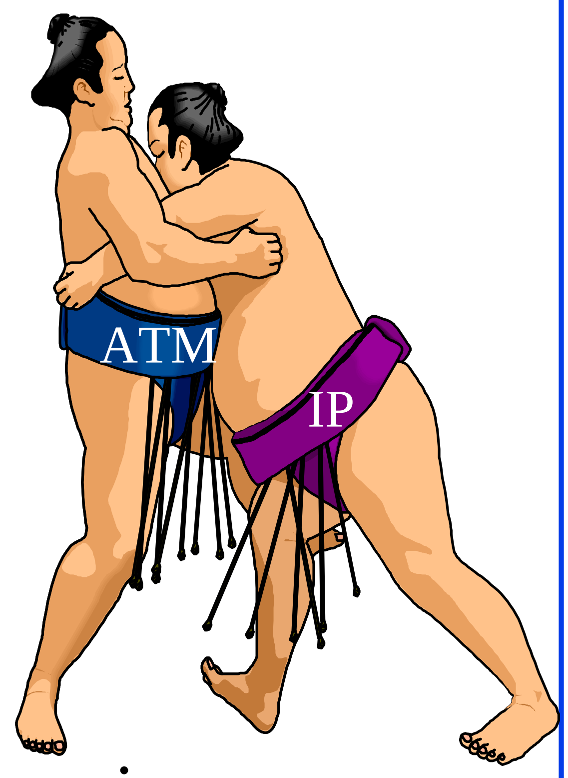
TCP/IP over ATM using ABR, UBR, and GFR Services and QoS over IP Issues

Raj Jain

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

- Why ATM?
- ABR: Binary and Explicit Feedback
- ABR Vs UBR
- TCP/IP over UBR
- TCP/IP over GFR
- QoS over IP: IntServ, DiffServ, MPLS

Why ATM?



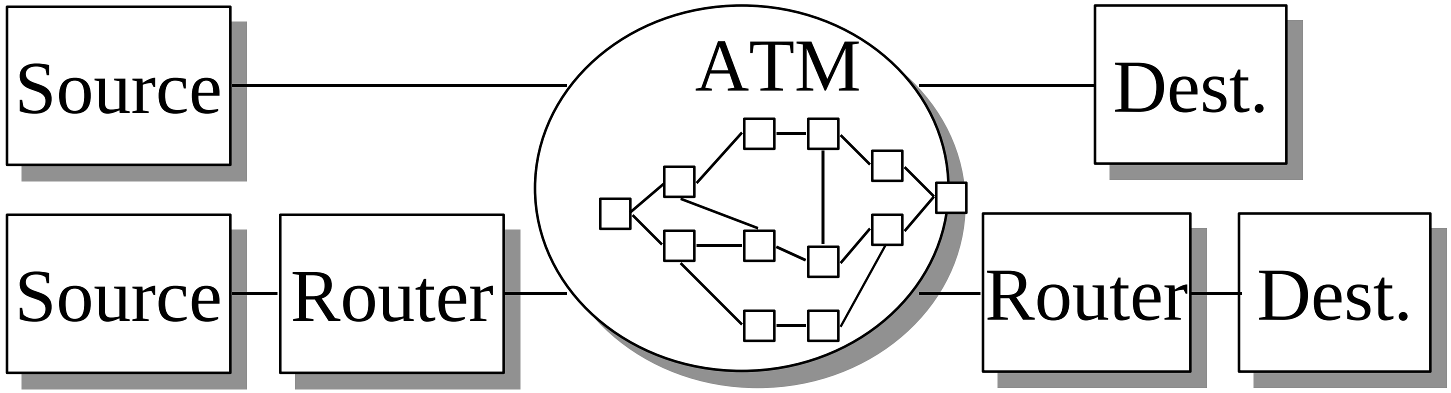
ATM vs IP: Key Distinctions

1. Traffic Management:
   Explicit Rate vs Loss based
2. Signaling: Coming to IP in the form of RSVP
3. QoS: PNNI routing, Service categories. Integrated/Differentiated services
4. Switching: Coming to IP as MPLS
5. Cells: Fixed size or small size is not important

Internet Protocols over ATM

- ATM Forum has designed ABR service for data
- UBR service provides no feedback or guarantees
- Internet Engineering Task Force (IETF) prefers UBR for TCP

ABR vs UBR



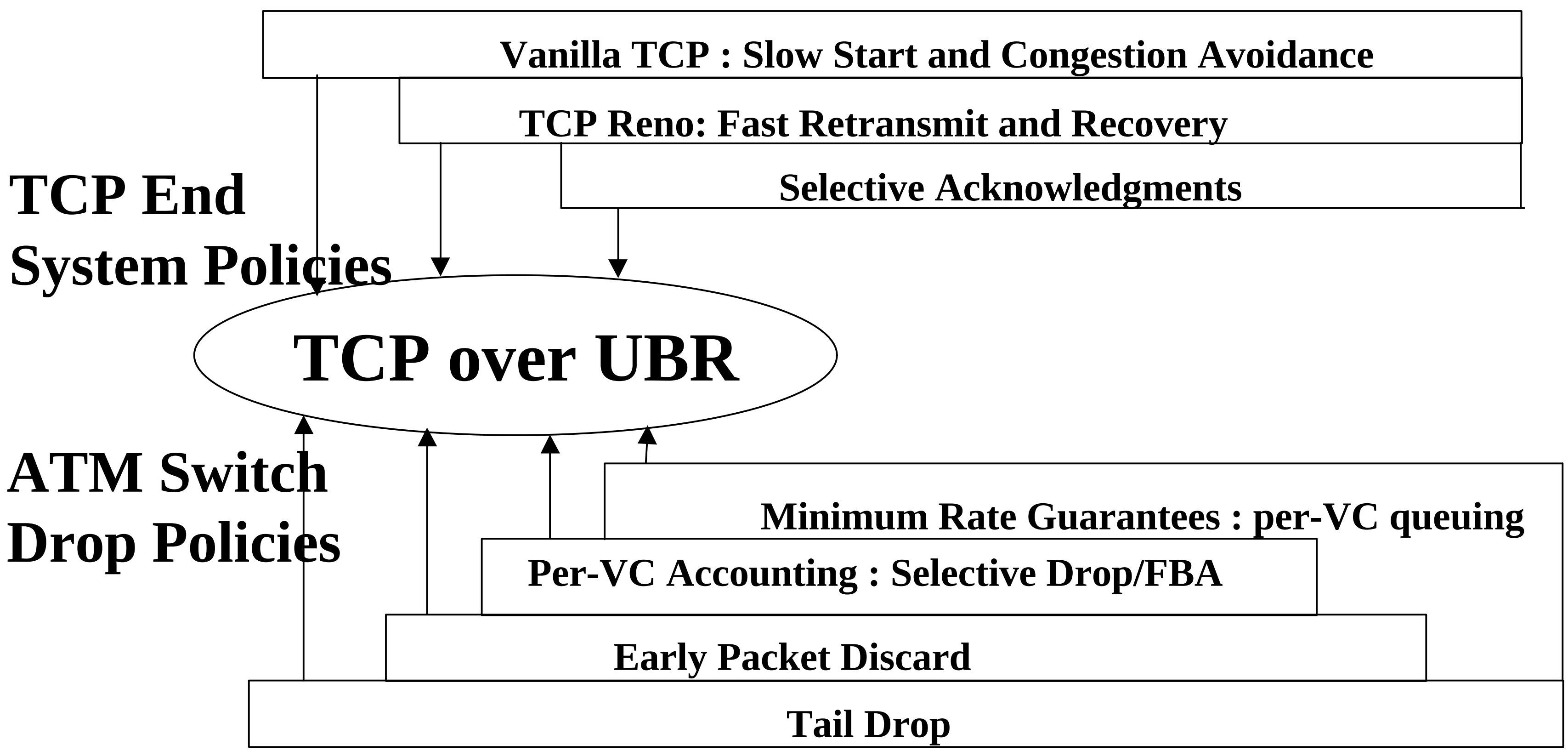
ABR

- Queue in the source
- Pushes congestion to edges
- If ATM not end-to-end: intelligent Q mgmt in routers
- Works for all protocols

UBR

- Queue in the network
- No backpressure
- Same end-to-end or backbone
- Works with TCP

Improving Performance of TCP over UBR



Policies

Switch Policies		No FRR	FRR	New Reno	SACK + New Reno
No EPD					
	Plain EPD				
EPD	Selective Drop				
	Fair Buffer Allocation				

End-System Policies

Policies: Results

- In LANs, switch improvements (PPD, EPD, SD, FBA) have more impact than end-system improvements (Slow start, FRR, New Reno, SACK). Different variations of increase/decrease have little impact due to small window sizes.
- In large bandwidth-delay networks, end-system improvements have more impact than switch-based improvements
- FRR hurts in large bandwidth-delay networks.

Policies (Continued)

- Fairness depends upon the switch drop policies and not on end-system policies
- In large bandwidth-delay networks:
  - SACK helps significantly
  - Switch-based improvements have relatively less impact than end-system improvements
  - Fairness is not affected by SACK
- In LANs:
  - Previously retransmitted holes may have to be retransmitted on a timeout
    $\Rightarrow$ SACK can hurt under extreme congestion.

Guaranteed Frame Rate (GFR)

- UBR with minimum cell rate (MCR) $\Rightarrow$ UBR+
- Frame based service
  - Complete frames are accepted or discarded in the switch
  - Traffic shaping is frame based. All cells of the frame have CLP =0 or CLP =1
  - All frames below MCR are given CLP =0 service. All frames above MCR are given best effort (CLP =1) service.

Guaranteed Rate Service

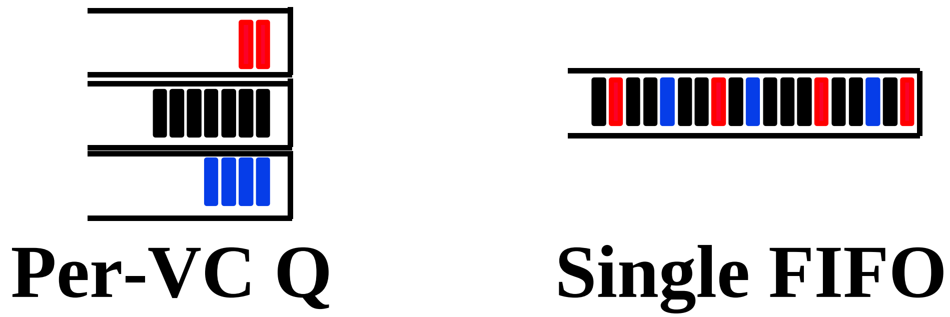
- Guaranteed Rate (GR): Reserve a small fraction of bandwidth for UBR class.

GR	GFR
per-class reservation	per-VC reservation
per-class scheduling	per-VC accounting/scheduling
No new signaling	Need new signaling
Can be done now	In TM4+

Guaranteed Rate: Results

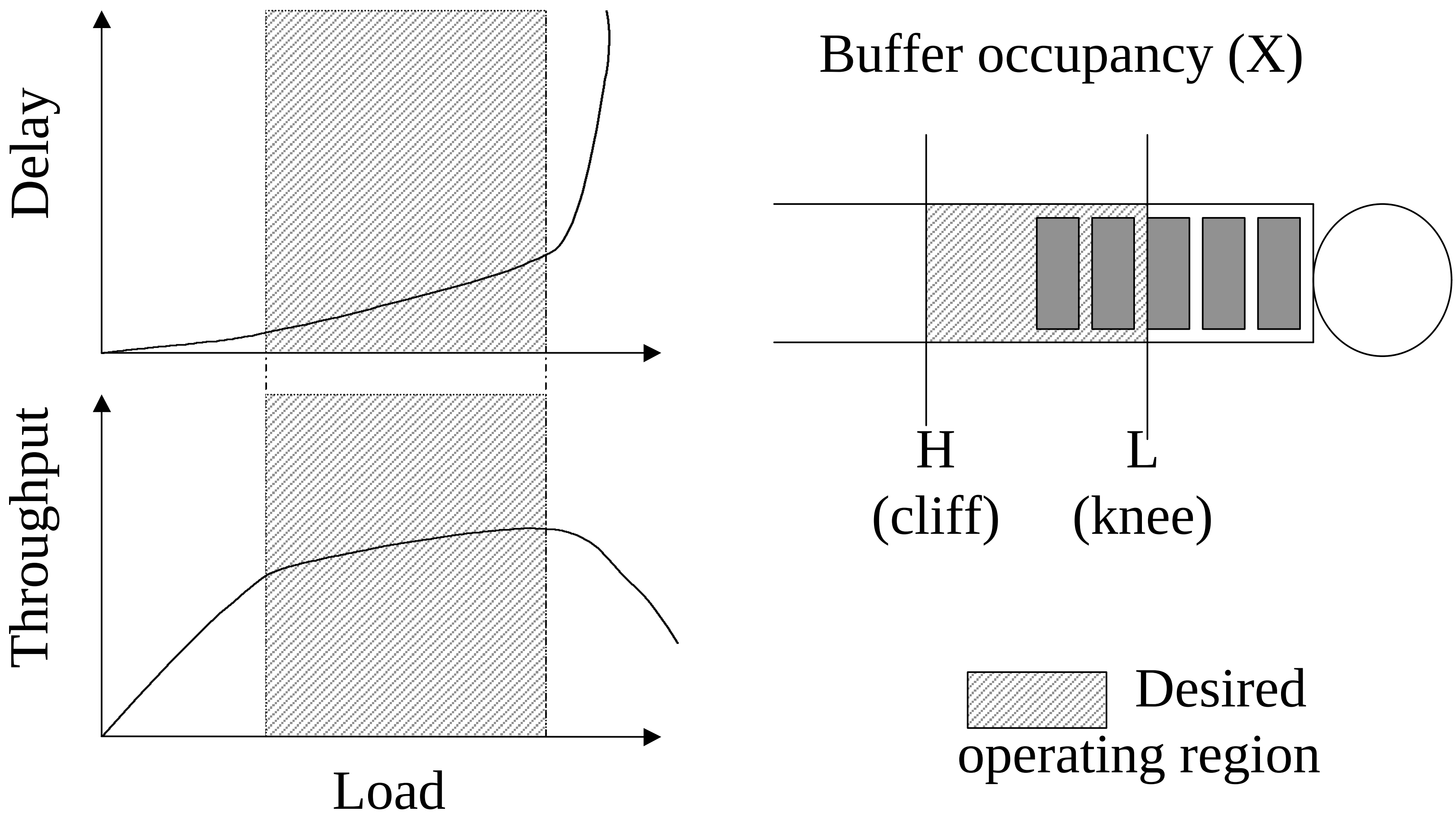
- Guaranteed rate is helpful in WANs.
- For WANs, the effect of reserving 10% bandwidth for UBR is more than that obtained by EPD, SD, or FBA
- For LANs, guaranteed rate is not so helpful. Drop policies are more important.

GFR: Results

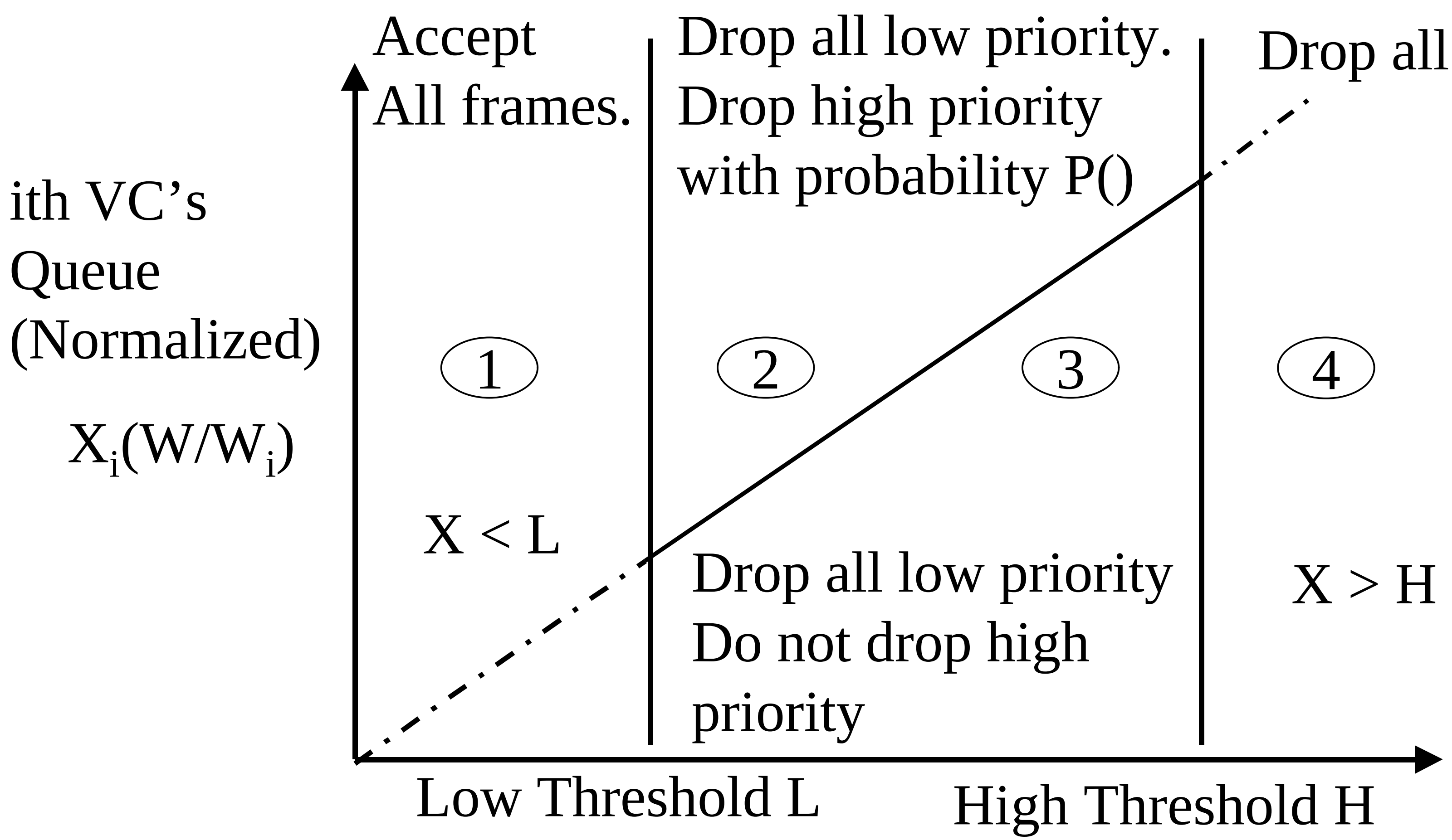


- Per-VC queuing and scheduling is sufficient for per-VC MCR.
- FBA and proper scheduling is sufficient for fair allocation of excess bandwidth
- Questions:
  - How and when can we provide MCR guarantee with FIFO?
  - What if each VC contains multiple TCP flows?

Differential Fair Buffer Allocation



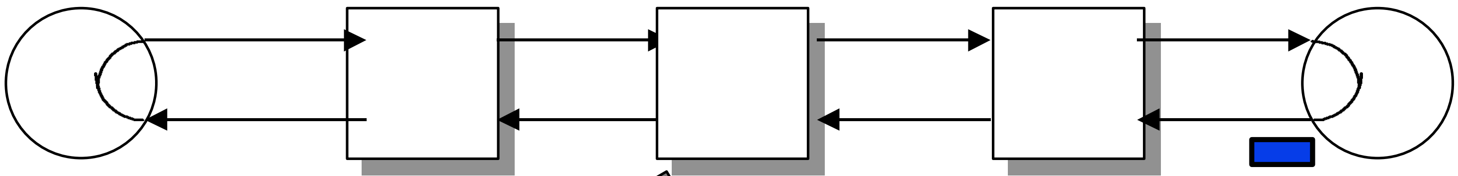
DFBA (contd.)



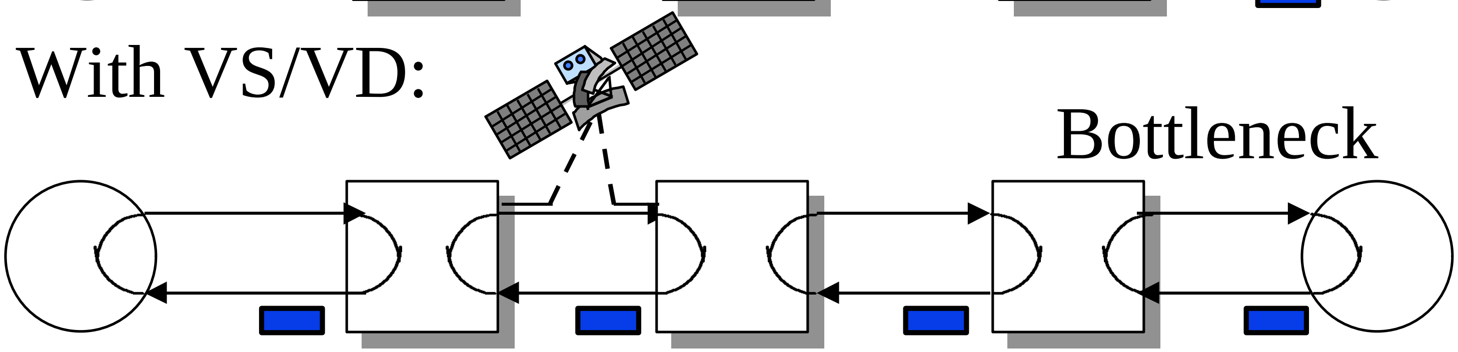
TCP Rate $D \propto \dfrac{MSS}{RTT \times \sqrt{P(drop)}}$

VS/VD

- Without Virtual Source/Virtual Destination:



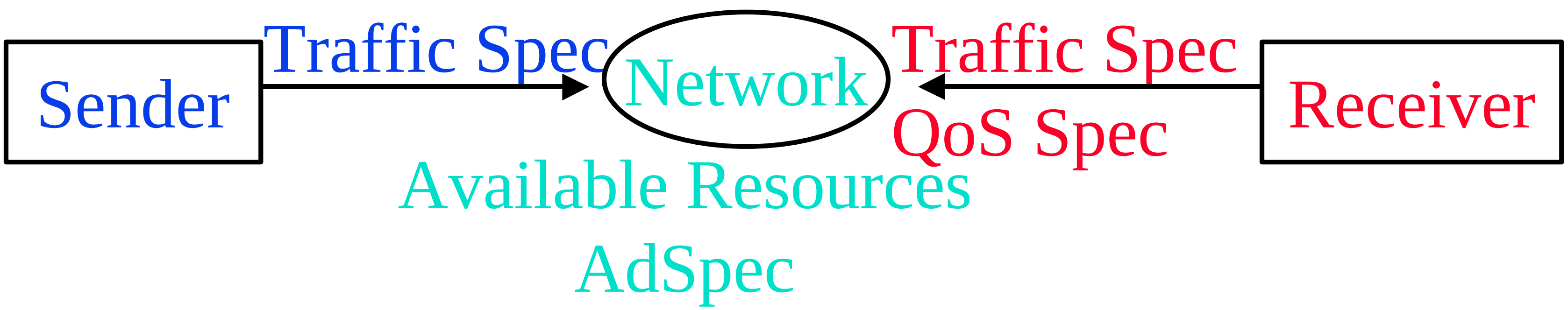
- With VS/VD:



- With VSVD, the buffering is proportional to the delay-bandwidth of the previous loop
  ⇒ Good for satellite networks

Integrated Services and RSVP

- 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: Like CBR or rt-VBR
  - Firm bound on data throughput and delay.
  - Is not always implementable, e.g., Shared Ethernet.
- Resource ReSerVation Protocol: Signaling protocol



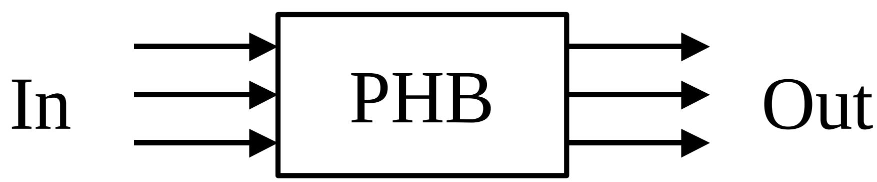
Problems with RSVP and Integrated Services

- Complexity: Packet classification, Scheduling
- Scalable in number of receivers per flow but
  Per-Flow State: O(n) $\Rightarrow$ Not scalable with # of flows.
  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.
- RSVP does not have negotiation and backtracking

Differentiated Services

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

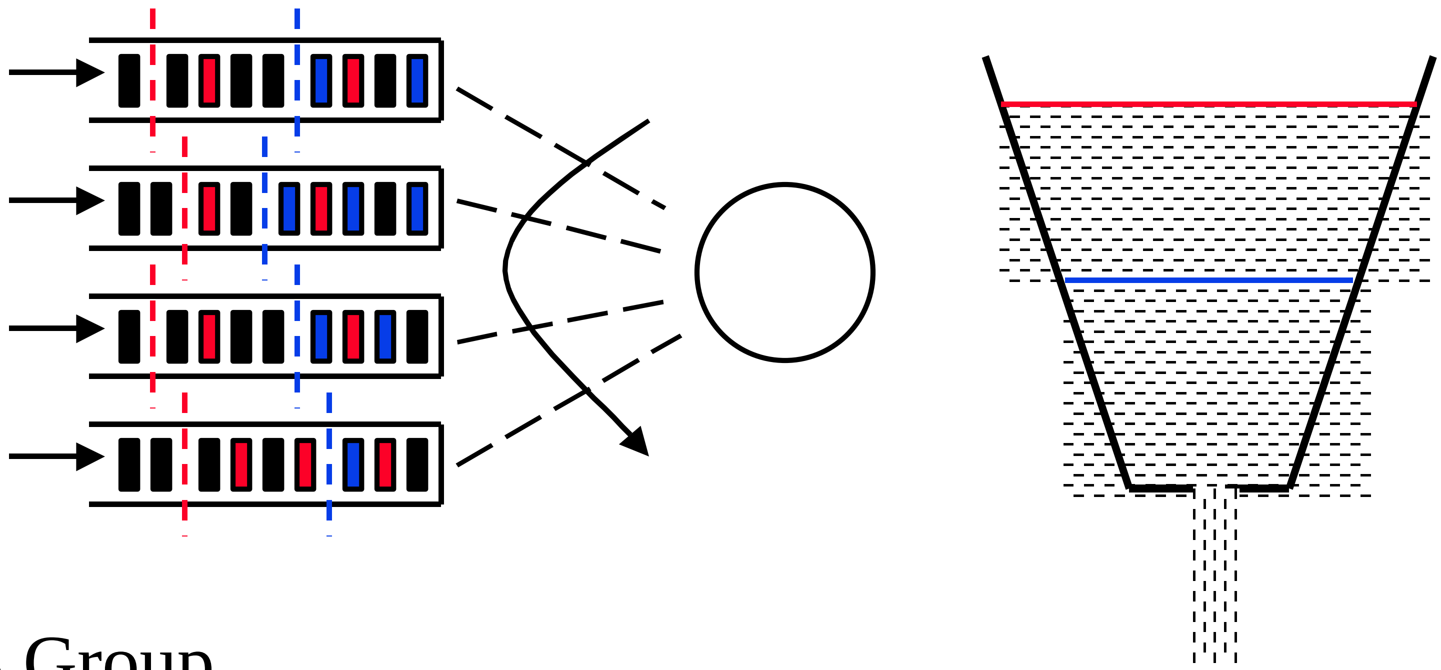
- IPv4: 3-bit precedence + 4-bit ToS
- Many vendors use IP precedence bits but the service varies ⇒ Need a standard ⇒ Differentiated Services
- DS working group formed February 1998
- Charter: Define ds byte (IPv4 ToS field)
- Per-Hop Behavior: Externally Observable Forwarding Behavior, e.g., x% of link bandwidth, or priority



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)

Assured Forwarding



- PHB <u>Group</u>
- Four Classes: Decreasing weights in WFR/WFQ
- Three drop preference per class
  (one rate and two bucket sizes)

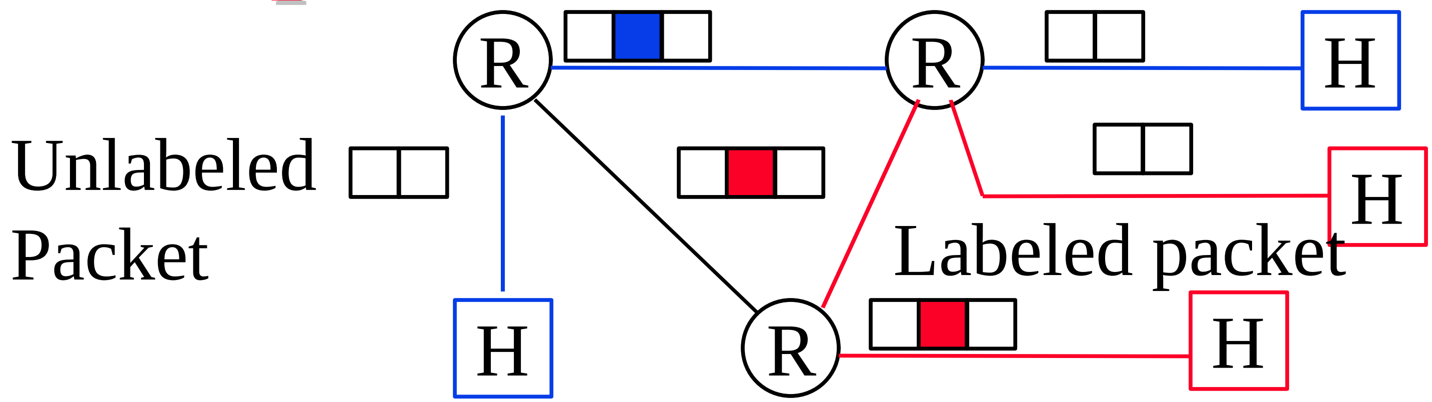
Problems with DiffServ

- per-hop $\Rightarrow$ Need at every hop
  One non-DiffServ hop can spoil all QoS
- End-to-end $\neq \Sigma$ per-Hop
  Designing end-to-end services with weighted guarantees at individual hops is difficult.
  Only EF will work.
- 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.

DiffServ Problems (Cont)

  - 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 $\Rightarrow$ Drop.
  Closed loop control $\Rightarrow$ Wait at source
  Data prefers waiting $\Rightarrow$ Feedback
- Guarantees $\Rightarrow$ Stability of paths
  $\Rightarrow$ Connections (hard or soft)
  Need route pinning or connections.

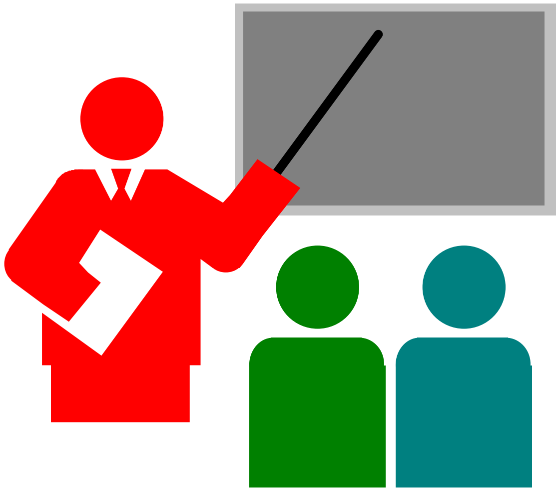
Multiprotocol Label Switching



- 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
  ⇒ Maximum throughput, Min delay, min loss
  ⇒ 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



Summary

- Traffic management distinguishes ATM from its competition
- Binary feedback too slow.
  ER switches better for high bandwidth-delay paths.
- ABR pushes congestion to edges.
  UBR+ may be OK for LANs but not for large bandwidth-delay paths.

Summary (Cont)

- Reserving a small fraction of bandwidth for the entire UBR class improves its performance considerably.
- It may be possible to do GFR with FIFO

Summary

- Multiple drop preferences does not help data (TCP) or Voice/Video
- Voice/video need multiple leaky bucket rates for layered/scalable coding.
- Need additivity or mathematical aggregatability. CBR (EF) should be the first step for IP.
- Excess allocation is useful with closed loop. Network/application dynamics ⇒ Need closed loop

Our Contributions and Papers

- All our contributions and papers are available on-line at http://www.cis.ohio-state.edu/~jain/
- See Recent Hot Papers for tutorials.