TCP 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/
Overview

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

Ref: For detailed studies, see
http://www.cis.ohio-state.edu/~jain/
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
**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 TCP over UBR

<table>
<thead>
<tr>
<th>Switch Policies</th>
<th>End-System Policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>No EPD</td>
<td>No FRR</td>
</tr>
<tr>
<td></td>
<td>FRR</td>
</tr>
<tr>
<td></td>
<td>New Reno</td>
</tr>
<tr>
<td></td>
<td>SACK + New Reno</td>
</tr>
<tr>
<td>EPD</td>
<td>Plain EPD</td>
</tr>
<tr>
<td></td>
<td>Selective Drop</td>
</tr>
<tr>
<td></td>
<td>Fair Buffer Allocation</td>
</tr>
</tbody>
</table>
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
  - 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 Rate Service

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

<table>
<thead>
<tr>
<th>GR</th>
<th>GFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>per-class reservation</td>
<td>per-VC reservation</td>
</tr>
<tr>
<td>per-class scheduling</td>
<td>per-VC accounting/scheduling</td>
</tr>
<tr>
<td>No new signaling</td>
<td>Need new signaling</td>
</tr>
<tr>
<td>Can be done now</td>
<td>In TM4+</td>
</tr>
</tbody>
</table>

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.
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

Buffer occupancy (X)

- H (cliff)
- L (knee)

Desired operating region

Throughput vs. Load

Delay vs. Load

The Ohio State University
DFBA (contd.)

ith VC’s Queue (Normalized): $X_i(W/W_i)$

- $X < L$
  - Accept All frames.
  - Drop all low priority.
  - Drop high priority with probability $P()$
  - Drop all

- $X > H$
  - Do not drop high priority
  - Total Queue $X$

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

The Ohio State University

Raj Jain
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

![Diagram of traffic and QoS specification flows between sender, network, and receiver]
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

- 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
  \[ \implies \text{Highest data priority (if priority queueing)} \]
Assured Forwarding

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

- per-hop ⇒ Need at every hop
  One non-DiffServ hop can spoil all QoS

- End-to-end ≠ \( \Sigma \) 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 ⇒ Drop.
  Closed loop control ⇒ Wait at source
  Data prefers waiting ⇒ Feedback
- Guarantees ⇒ Stability of paths
  ⇒ Connections (hard or soft)
  Need route pinning or connections.
- 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
Traffic management distinguishes ATM from its competition

ABR pushes congestion to edges. UBR+ may be OK for LANs but not for large bandwidth-delay paths.

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.