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

Raj Jain
The Ohio State University

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

state.edu/~jain/talks/qos9906.htm
Overview

- 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
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
- ATM also has QoS-based routing (PNNI)
ATM QoS

Too much too soon
ATM QoS: Issues

- Can’t easily aggregate QoS: VP = \( \Sigma \) VCs
- Can’t easily specify QoS: What is the CDV required for a movie?
- Signaling too complex ⇒ 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

**Diagram:**
- Sender
- Network
- Available Resources
- AdSpec
- Receiver
- Traffic Spec
- QoS Spec
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: $O(n) \Rightarrow$ 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.
  $\Rightarrow$ 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.
- RSVP does not have negotiation and backtracking
- 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.
Differentiated Services

<table>
<thead>
<tr>
<th>Ver</th>
<th>Hdr Len</th>
<th>Precedence</th>
<th>ToS</th>
<th>Unused</th>
<th>Tot Len</th>
</tr>
</thead>
<tbody>
<tr>
<td>4b</td>
<td>4b</td>
<td>3b</td>
<td>4b</td>
<td>1b</td>
<td>16b</td>
</tr>
</tbody>
</table>

- 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)
Service

- 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
  ⇒ 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:

<table>
<thead>
<tr>
<th>Drop Prec.</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>010 000</td>
<td>011 000</td>
<td>100 000</td>
<td>101 000</td>
</tr>
<tr>
<td>Medium</td>
<td>010 010</td>
<td>011 010</td>
<td>100 010</td>
<td>101 010</td>
</tr>
<tr>
<td>High</td>
<td>010 100</td>
<td>011 100</td>
<td>100 100</td>
<td>101 100</td>
</tr>
</tbody>
</table>

- 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 better performance for TCP flows when there is considerable unused bandwidth.

Reason: TCP does not get any share of excess bandwidth in presence of UDP.

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 ⇒ Controls buffer allocation  
  ⇒ Indirectly affects bandwidth allocation
  - Depends upon the arrival pattern  
    ⇒ Random ⇒ Not Reliable
- Given a limited number of PHB’s, it is better to have more classes than more DPs
DiffServ Problems (Cont)

- DiffServ is unidirectional $\implies$ No receiver control
- Modified DS field $\implies$ 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

- Label = Circuit number = VC Id
- Ingress router/host puts a label. Exit router strips it off.
- Switches switch packets based on labels. Do not need to look inside ⇒ Fast.

The Ohio State University
Traffic Engineering Objectives

- User’s Performance Optimization
  ⇒ Maximum throughput, Min delay, min loss, min delay variation

- Efficient resource allocation for the provider
  ⇒ Efficient Utilization of all links
  ⇒ Load Balancing on parallel paths
  ⇒ Minimize buffer utilization
  - Current routing protocols (e.g., RIP and OSPF) find the shortest path (may be over-utilized).

- QoS Guarantee: Selecting paths that can meet QoS
- Enforce Service Level agreements
- Enforce policies: Constraint based routing ⇒ QoS
MPLS Mechanisms for TE

- Signaling, Admission Control, Routing
- Explicit routing of LSPs
- Constrained based routing of LSPs
  Allows both Traffic constraints and Resource Constraints (Resource Attributes)
- Hierarchical division of the problem (Label Stacks)
- Traffic trunks allow aggregation and disaggregation
  (Shortest path routing allows only aggregation)
Traffic Trunks

- Trunk: Aggregation of flows of same class on same LSP
- Trunks are routable
  - ⇒ LSP through which trunk passes can be changed
- Class ⇒ Queue, LSP ⇒ Next hop
  - Class can be coded in Exp or Label field. Assume Exp.
Trunks vs LSPs

Tour Group

Flights = LSP
Tour Groups = Trunks
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

<table>
<thead>
<tr>
<th>DL</th>
<th>Label</th>
<th>Exp</th>
<th>SI</th>
<th>TTL</th>
<th>IP</th>
<th>TCP</th>
</tr>
</thead>
</table>

Flows, Trunk

LSP

Link
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

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

<table>
<thead>
<tr>
<th>Dest Addr</th>
<th>Src Addr</th>
<th>Tag Prot ID</th>
<th>Pri</th>
<th>CFI</th>
<th>VLAN ID</th>
</tr>
</thead>
</table>

802.1Q header

- 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
QoS Debate Issues

- Massive Bandwidth vs Managed Bandwidth
- Per-Flow vs Aggregate
- Source-Controlled vs Receiver Controlled
- Soft State vs Hard State
- Path based vs Access based
- Quantitative vs Qualitative
- Absolute vs Relative
- End-to-end vs Per-hop
- Static vs Feedback-based
- One-way multicast vs n-way multicast
- Homogeneous multicast vs heterogeneous multicast
- Single vs multiple bottlenecks: Scheduling
## Comparison of QoS Approaches

<table>
<thead>
<tr>
<th>Issue</th>
<th>ATM</th>
<th>IntServ</th>
<th>DiffServ</th>
<th>MPLS</th>
<th>IEEE 802.3D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massive Bandwidth vs Managed Bandwidth</td>
<td>Managed</td>
<td>Managed</td>
<td>Massive</td>
<td>Managed</td>
<td>Massive</td>
</tr>
<tr>
<td>Per-Flow vs Aggregate</td>
<td>Both</td>
<td>Per-flow</td>
<td>Aggregate</td>
<td>Both</td>
<td>Aggregate</td>
</tr>
<tr>
<td>Source-Controlled vs Receiver Controlled</td>
<td>Unicast</td>
<td>Receiver</td>
<td>Ingress</td>
<td>Both</td>
<td>Source</td>
</tr>
<tr>
<td>Source, Multicast both</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soft State vs Hard State</td>
<td>Hard</td>
<td>Soft</td>
<td>None</td>
<td>Hard</td>
<td>Hard</td>
</tr>
<tr>
<td>Path based vs Access based</td>
<td>Path</td>
<td>Path</td>
<td>Access</td>
<td>Path</td>
<td>Access</td>
</tr>
<tr>
<td>Quantitative vs Qualitative</td>
<td>Quantitative</td>
<td>Quantitative</td>
<td>Mostly qualitative</td>
<td>Both</td>
<td>Qualitative</td>
</tr>
<tr>
<td>Absolute vs Relative</td>
<td>Absolute</td>
<td>Absolute</td>
<td>Mostly Relative</td>
<td>Absolute plus relative</td>
<td>Relative</td>
</tr>
</tbody>
</table>

The Ohio State University

Raj Jain
### Comparison (Cont)

<table>
<thead>
<tr>
<th>Issue</th>
<th>ATM</th>
<th>IntServ</th>
<th>DiffServ</th>
<th>MPLS</th>
<th>IEEE 802.3D</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-to-end vs Per-hop</td>
<td>e-e</td>
<td>e-e</td>
<td>Per-hop</td>
<td>e-e</td>
<td>Per-hop</td>
</tr>
<tr>
<td>Static vs Feedback-based</td>
<td>Both</td>
<td>Static</td>
<td>Static</td>
<td>Static</td>
<td>Static</td>
</tr>
<tr>
<td>One-way multicast vs n-way multicast</td>
<td>Only one-way</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homogeneous multicast vs heterogeneous multicast</td>
<td>Homogeneous</td>
<td>Heterogeneous</td>
<td>N/A</td>
<td>Homogeneous</td>
<td>N/A</td>
</tr>
<tr>
<td>Single vs multiple bottlenecks: Scheduling</td>
<td>Multiple bottleneck</td>
<td>Multiple</td>
<td>Multiple</td>
<td>Multiple</td>
<td></td>
</tr>
</tbody>
</table>
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!