The problems in trying to create a QoS Internet

Scott Bradner

Harvard University
sob@harvard.edu

Internet Economics

- current ISP offering are all “best effort”
- competition is based on lowest cost service
- there is no technology to ensure “better service”
  beyond than answering the phone faster
- only option is less over subscription
- very hard to see how ISPs can be profitable
  in the long run

Other Network Economics (US)

- telephone network - $200 billion / 155 million lines
  - $100 billion long distance
- broadcast TV - $25 billion / 96 million homes
- video rental - $9 billion
- cable TV - $3.9 billion / 59 million homes
- pay per view - $660 million
- EDI - $122 million
- ISP - $3.3 billion

Modified Internet Economics

- very large potential revenue if the Internet could provide differential QoS
- users would pay more for better service
- migrate some services to the Internet
  - even a small migration of long distance could have a major impact on ISP revenue
- are guarantees required?
  - or are probabilities sufficient?

QoS Targets

- capacity
- loss rate
- latency
- data flow rate
- predictable behavior for users
  McDonnalds vs. FedEx

Quality of Service, Background

- big call for QoS
- raison d'être for ATM
- push behind RSVP
- confusion over meaning, type and need

Quality of Service, What Is It?

- the ability to define or predict aspects of the performance of systems on a network
- long-time "glass house" requirement
  - SNA is seen as having lots of QoS controls
  - connection-oriented protocol
- one of the original goals for the Internet Protocols
  - "type of service" - differ in speed, latency & reliability
  - datagram protocol (for robustness)

Where is QoS Needed?

- where there are constrained resources
  - lines
  - interconnect devices
  - servers
- if you have enough resources, QoS controls generally not required
- except if consistent response wanted
  - may need to slow things down

QoS Types

- predictive
  - architect network based on observed loads
  - can also police input loads
- flow based
  - reserve capacity during the execution of an application
  - keep track of reservation in each network device in path
  - conflict with datagram mode of IP
- non flow based
  - mark packets to indicate class
  - process packets differently in network based on marking

Predictive QoS

- QoS in most current datagram networks
- “just” make network “big” enough
- reasonable on a LAN or campus network
- hard to do for WAN
- tends to provide cycles of quality
  - over build for need
  - need catches up and passes capacity
  - over build for new need

Throw Bandwidth at Problem

- with “enough” bandwidth QoS can be easy
  - enough means much more than peaks
  - e.g., gigabit Ethernet for 1 video stream
- still might have to sequence data onto link
  - if bursty traffic

Flow Based QoS

- per flow state kept in network
- ATM & RSVP QoS are the per flow type
- scaling issues
- authentication issues
- accounting issues

Flow Based QoS

- ATM QoS
- IP-based QoS
- mixed

ATM QoS

- basic QoS is to control:
  - absolute cell latency from source to destination
  - variation in cell latency
- different requirements for broadcast vs. interactive
- tension between low variation and reliable data
  - low variation means small buffers
  - reliable LAN data means large buffers
  - can make sure that specific VCs have small buffers and high priority to ensure low latency variation
- problem since ATM QoS expects ATM end-to-end
  - real world has Ethernet or Token Ring at ends

Integrated Services (Int-Serv)

- architecture for supporting real-time applications over the Internet Protocols and the Internet
- guaranteed delay bounds
  absolute upper bound of delay
- link sharing
  set maximum shares of a link
- predictive real-time service
  stable delay
- overview - Informational RFC 1633

Int-Serv, contd.

- basic parts
  admission control - determines if new flow can be added
  classifier - determines flow for incoming packet
  packet scheduler - queues packets for transmission
    also requires an estimator for outgoing packet stream
    can use Weighted Fair Queuing (WFQ)
- not just traffic prioritization on a link

Int-Serv, contd.

- priority by itself is not enough
  - if too much high-priority traffic, prioritization does not help
  - separate request process, not accepted if it would overload
- requires flow-specific state in routers
  - change in basic Internet model
  - use soft state - can change on path change
    - vs. hard state - set at start, teardown at end
- may require request & flow authentication

RSVP

- Resource ReServation Protocol (RSVP)
- implementation of INTSRV reservation process
- can be used to set aside resources for a specific application along a communications path
- can transfer the requests to a new path if rerouted
- simplex (one direction per reservation)
- receiver-oriented
- may make use of QoS-active links
  - like ATM

RSVP - Process

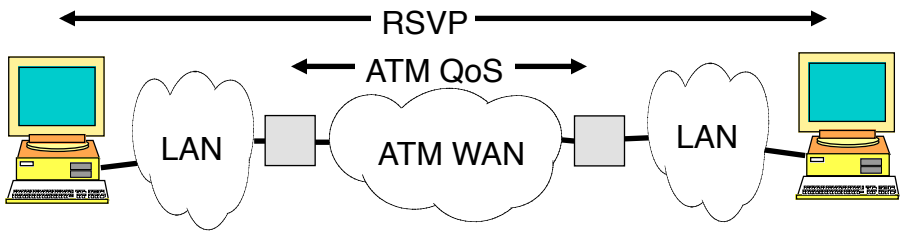
- sender transmits **path** messages to receiver
  - routers store path state
  - path message may also include
    - *sender template* - what do the packets "look" like
    - *tspec* - upper bound on traffic sender will send
- receiver sends **resv** messages back to sender
  - routers forward based on path state
  - resv messages include
    - *flowspec* - define a requested QoS
    - *filter_spec* - define specific packets for flowspec
    - *policy_data* - info for policy decision on acceptance
    - *integrity* - originating node authentication

RSVP - Process, contd.

- using admission control, router
  - will accept reservation request if enough capacity
    - record reservation and forward **`resv`** to next-hop
  - if not - send **`resverr`** to previous hop
- state refreshed periodically with new messages
  - entry removed on timeout
- periodic refresh deals with reroute

Mixed QoS

- since only sure end-to-end technology is IP
  (for now)
- use IP signaling (like RSVP) to control link-level QoS (like ATM) when present



Flow Based QoS Issues

- scaling issues
- authorization issues
- accounting issues
- advanced reservations *very* hard
- good for long flows (video, audio, large file transfers)
  flow setup cost must be low when averaged over flow length

Flow Lengths in the Internet

from cic nets' Chicago hub

IP Flow Switching Cache, 16384 active flows, 0 inactive
  132159644 added, 124468367 replaced, 4892577 timed out, 2782316 invalidated
  statistics cleared 270640 seconds ago

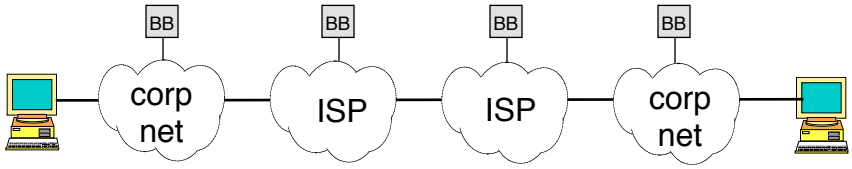
Protocol	Total Flows	Flows /Sec	Packets /Flow	Bytes /Pkt	Packets /Sec	Active(Sec) /Flow	Idle(Sec) /Flow
TCP-Telnet	5222464	19.2	40	89	785.3	32.9	17.3
TCP-FTP	2087345	7.7	6	87	47.9	7.3	22.7
TCP-FTPD	1275958	4.7	95	390	449.5	21.9	23.6
TCP-WWW	83916123	310.0	9	304	2944.5	5.4	20.9
TCP-SMTP	14106833	52.1	8	173	448.9	6.4	21.6
TCP-X	94849	0.3	81	176	28.6	24.1	17.8
TCP-other	16095661	59.4	38	274	2290.8	20.9	21.5
UDP-TFTP	339	0.0	1	207	0.0	2.3	21.0
UDP-other	5059444	18.6	11	217	208.4	9.4	26.0
ICMP	4201689	15.5	2	83	46.0	5.2	26.8
IGMP	39809	0.1	30	398	4.4	48.2	29.4
IPINIP	9431	0.0	1808	254	63.0	147.1	18.6
GRE	32811	0.1	594	204	72.0	62.1	18.8
IP-other	909	0.0	3	223	0.0	1.2	31.8
Total:	132143665	488.2	15	260	7389.7	0.0	0.0

Non Flow Based QoS

- packet headers are "marked" at edge of network
  - precedence bits most common place to mark
- one or more bits used
  - two (priority and best effort) or more levels
- different mechanisms proposed
  - drop priority, queue selector - WFQ on queues
- contract with ISP, contract between ISPs
  - a problem if too much traffic for destination
  - unless admission control (marker) is complex
- new (unproven) ideas
- can create N predictive Vnets on same Pnet

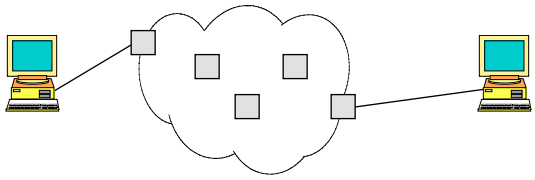
Virtual Leased Line

- proposal from Van Jacobson
- produce a virtual network that acts like a physical leased line network
- simulate the router “leaky bucket” operation
- includes complex “bandwidth broker” logic
- easy to explain to a customer



Edge Shaping

- proposal from Dave Clark & John Wroclawski
- simple queuing process in routers
- differential drop rates based on “in/out” bit
  in or out of flowspec
- edge devices set bit using complex mechanisms
- aim is to work with TCP flow control



Simple Differential Service

- proposal from Paul Ferguson
- mark packets with a "delay indicator" &/or "drop preference"
- hint to router on relative proprity of the packet
- router mechanisms not included

What Should IETF do?

- standardize mechanisms?
- standardize bit semantics?
- standardize what bits change?
- standardize services?