Does reality matter?: QoS & ISPs

Scott Bradner
Harvard University

In the Beginning

- in the beginning (and now)
- there was (is) philosophy
  - or is that religion?
- smart network vs. smart edges
- centralized vs. distributed
- circuits vs. datagrams
- redundancy vs. reliability for reliability

So What Happened?

- telco world went with circuits
  - X.25, frame relay, ISDN, ATM
- Internet went with datagrams
  - ARPANET, NSFNET, Internet
- telco world went with smart network
  - SS7, dumb edges, applications in telco switches
- Internet went with dumb net
  - soft (if any) state in net, smart edges, applications in edges

Innovation?

- telco world
  - innovation = *69
- Internet
  - innovation = www
- telco world
  - standards to preserve power status quo
- Internet
  - standards to create technology

Implications of Circuit vs Packet

- paths through network are not stable
  - change based on
    - link failure
    - traffic engineering
    - routing instability
    - link utilization (someday)
- impacts QoS
  - hard to reserve resources
  - unpredictable QoS
  - IBM: *"can not build corporate network using TCP/IP"*

s&rbn - 5

Quality of Service (QoS)

- the ability to define or predict the performance of systems on a network
  - note: predictable may not mean "best"
- unfair allocation of resources under congestion conditions
  - Bill pays to get Fred's traffic dropped

$$$

- long-time SNA feature

s&rbn - 6

Is QoS Important in the Internet?

- pundits want QoS, some purists are not sure
  - do you want to block an emergency phone call?
- is service definition a point?
  - or a curve?
    - remember cell phones
- QoS targets
  - telephone bypass (e.g. phone over cable modems)
  - IP voice trunking

Interactive Applications

- max latency determined by some external constraint
  - e.g. human systems
    - max RTT for voice interaction 300 - 400 msec
    - otherwise talk over each other
- data that is too late is useless
- but significant % loss still works

Now What?

- ISP has decided that telco bypass = $$$
- IP seems to be the answer
  - it's the answer to everything else, why not this?
- for IP trunking
  - throw bandwith at problem
    - or use "good" ISP & high speed links
    - much in use for international regulatory by-pass
- but customer phones seem different
  - lets look at using IP
    - assumption is that QoS controls are needed

IP & QoS

- original goal in IP - TOS bits - RFC 791
  - *provides an indication of the abstract parameters of the quality of service desired*
  - *guide the selection of the actual service parameters when transmitting a datagram through a particular network*
  - intended to be used only within a single network
- *expected to be used to control ... routing and queuing algorithms* (RFC1122)
- *precedence is a scheme for allocating resources in a network based on the importance of different traffic flows* (RFC 1812)

NETWORLD+INTEROP

What Happened to IP QoS

- never quite focused on the issue
- general answer - more bandwidth
- return was not worth the added complexity
  - e.g. TOS routing removed from OSPF
- but if you are determined to get IP QoS

QoS Types

- predictive
  - architect network based on observed loads
  - can also police input loads
- flow based
  - reserve bandwidth through network for an execution of an application
  - keep track of reservation in each network device in path
- non flow based
  - mark packets to indicate class
  - process 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
- no actual guarantees
- 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

s&rbn - 13

Flow Based QoS

- traditional telco answer
- per flow reservations
- per flow guarantees
- per flow state kept in network
  - e.g. X.25, frame relay, ATM
- has scaling issues
- IETF per-flow QoS work
  - inteserv - link level mechanisms
  - RSVP - signaling

s&rbn - 14

NETWORLD +INTEROP

ATM QoS

- set up virtual circuit across network
  - defined QoS for each VC
- basic ATM QoS is designed to control:
  - absolute cell latency from source to destination
  - variation in cell latency
- once thought that you could set up VC for each datagram
  - but performance not there
- could use VC per phone call
  - if ATM were end2end

IETF Integrated Services

- assume desire to use the Internet as common infrastructure for real-time and non-real-time communication
- two defined services
  - guaranteed
  - controlled-load

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
- may make use of QoS-active links
  - like ATM if there
- ATM imitation

s&rbn - 17

Flow Based QoS Issues

- scaling issues - per flow state an issue
- authorization (policy) issues - who says "OK"
- accounting issues - how to bill user
- security issues - theft / denial of service
- advanced reservations *very* hard
- good for long flows (video, audio, large file transfers, VPNs)
  - flow setup cost must be low when averaged over flow length
- many mice on the Internet

s&rbn - 18

Flow Lengths in the Internet

from cic nets' Chicago hub

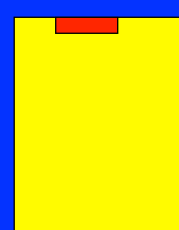
IP Flow Switching Cache, 16384 active flows, 0 inactive
132159644 added, 124468367 replaced, 4892577 timed out, 2782316 invalidated
statistics cleared 270640 seconds ago

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

s&rbn - 19

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
- new (unproven) ideas
- creates N predictive Vnets on same Pnet



s&rbn - 20

NETWORLD+INTEROP

Non Flow Based QoS, contd.

- 1st model = "sender pays"
  - "receiver pays" may come later
- can use long or short term QoS contracts with ISP
  - dynamic requests for more bandwidth
- better scaling than per flow QoS
- easier authentication, authorization and accounting
- still much research needed
- hard (very hard) to get actual guarantees

Non Flow Based Qos in the IETF

- IETF Differentiated Services working group
- does not replace intserv /RSVP
- to define class-based QoS
  - replace earlier definition of use of TOS byte
- 1st define behaviors not services
  - now thinking about services
- will look at traffic shapers & packet markers

IETF Diffserv WG

- rename IP TOS Byte to "DS Field"
- components
  - mark bits in DS Field at network "edge"
  - routers in net use markings to determine packet treatment
  - conditioning marked packets at network boundaries
- deals with flow aggregates
- DS Field may change in flight
  - some disagreement - what about end-to-end?
- note! - diffserv not guaranteed service
  - does not know "destination"

Basic PHBs

- base difserv RFC includes precedence field computability - RFC 2474
- PHB = 000000 default (best effort)
- PHB = xxx000 ordered priority handling
  - backward compatible with precedence bits

Expedited Forwarding (EF)

- one PHB
- strict policing at edges
  - to ensure no overload in network
- produces a guaranteed service
  - assuming correct admission control
- requires system to coordinate edge policing
  - proposal for a “Bandwidth Broker”
- departure rate of traffic must equal or exceed a configurable rate

Assured Forwarding Group (AF)

- set of PHBs
  - 4 sets of 3 PHBs
  - organized as 4 queues, each with 3 levels of drop precedence
    - traffic must be forwarded based on precedence - not absolute priority
  - no specific ordering between classes
- can be used to provide frame-relay like services
- assured rather than guaranteed
- depends on edge policing & marking
  - can remark drop precedence in net

Reality

- qos policy
  - when to give a busy signal
- is it end-to-end?
  - what does the host have to say about it?
- $$$$
  - not just best effort
    - customers & peers
  - how should ISPs do settlements?
- is added complexity worth it?

Complexity not Worth It

- is adding bandwidth all that's needed?
- Andrew Odlyzko of AT&T Labs
  - may be cheaper to just throw bandwidth at QoS problem
  - 1 - only a few points of congestion
  - 2 - 80% of data com costs non-transmission
  - 3 - adding QoS complexity will add to other costs
    - labor, management & billing systems etc
  - 4 - local part of data com dominate overall cost
  - 5 - cost of transmission coming down
    - Fortune reports - 99.8 Tbps capacity by 2001 = glut
  - upgrade congested points - cheaper than QoS complexity

So Does QoS Make Sense for ISPs?

- no
  - in a well engineered core
- yes
  - for customer tail circuits
- jury still out
  - between ISPs
  - to cash in on telco $$$$$$$
  - server support
- still magic
  - control systems