

Quality of Service

Outline

Realtime Applications
Integrated Services
Differentiated Services

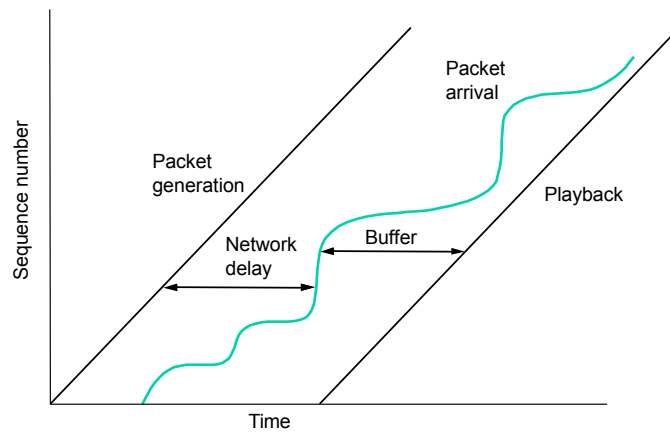
Realtime Applications

- Require “deliver on time” assurances
 - must come from *inside* the network

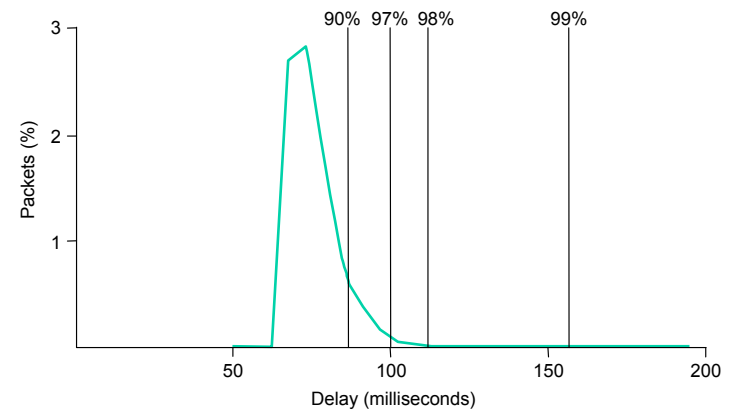


- Example application (audio)
 - sample voice once every 125us
 - each sample has a *playback time*
 - packets experience variable delay in network
 - add constant factor to playback time: *playback point*

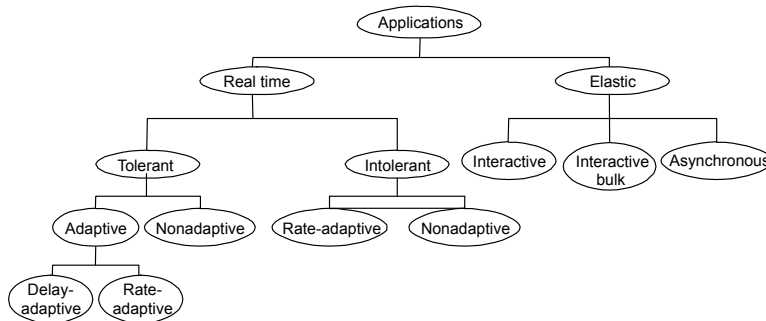
Playback Buffer



Example Distribution of Delays



Taxonomy



Integrated Services

- Service Classes
 - guaranteed
 - controlled-load
- Mechanisms
 - signalling protocol
 - admission control
 - policing
 - packet scheduling

Flowspec

- ***Rspec***: describes service requested from network
 - controlled-load: none
 - guaranteed: delay target
- ***Tspec***: describes flow's traffic characteristics
 - average bandwidth + burstiness: *token bucket* filter
 - token rate r
 - bucket depth B
 - must have a token to send a byte
 - must have n tokens to send n bytes
 - start with no tokens
 - accumulate tokens at rate of r per second
 - can accumulate no more than B tokens

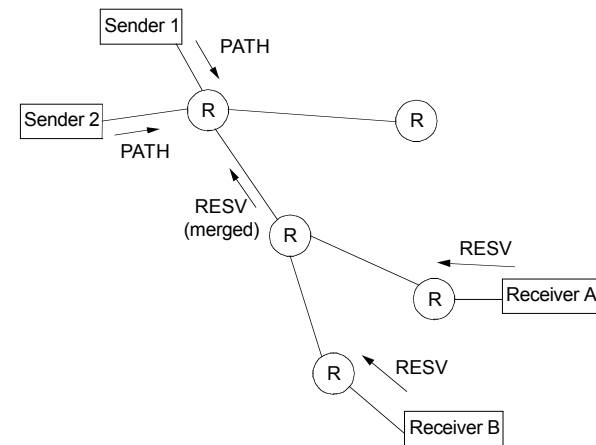
Per-Router Mechanisms

- Admission Control
 - decide if a new flow can be supported
 - answer depends on service class
 - not the same as *policing*
- Packet Processing
 - classification: associate each packet with the appropriate reservation
 - scheduling: manage queues so each packet receives the requested service

Reservation Protocol

- Called signaling in ATM
- Proposed Internet standard: RSVP
- Consistent with robustness of today's connectionless model
- Uses soft state (refresh periodically)
- Designed to support multicast
- Receiver-oriented
- Two messages: PATH and RESV
- Source transmits PATH messages every 30 seconds
- Destination responds with RESV message
- Merge requirements in case of multicast
- Can specify number of speakers

RSVP Example



RSVP versus ATM (Q.2931)

- RSVP
 - receiver generates reservation
 - soft state (refresh/timeout)
 - separate from route establishment
 - QoS can change dynamically
 - receiver heterogeneity
- ATM
 - sender generates connection request
 - hard state (explicit delete)
 - concurrent with route establishment
 - QoS is static for life of connection
 - uniform QoS to all receivers

Differentiated Services

- Problem with IntServ: scalability
- Idea: segregate packets into a small number of classes
 - e.g., premium vs best-effort
- Packets marked according to class at edge of network
- Core routers implement some per-hop-behavior (PHB)
- Example: Expedited Forwarding (EF)
 - rate-limit EF packets at the edges
 - PHB implemented with class-based priority queues or WFQ

DiffServ (cont)

- Assured Forwarding (AF)
 - customers sign service agreements with ISPs
 - edge routers mark packets as being “in” or “out” of profile
 - core routers run RIO: RED with in/out

