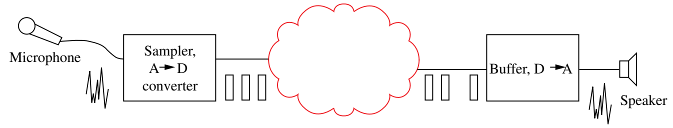
Quality of Service

- Outline
  - Realtime Applications
  - Integrated Services
  - Differentiated Services

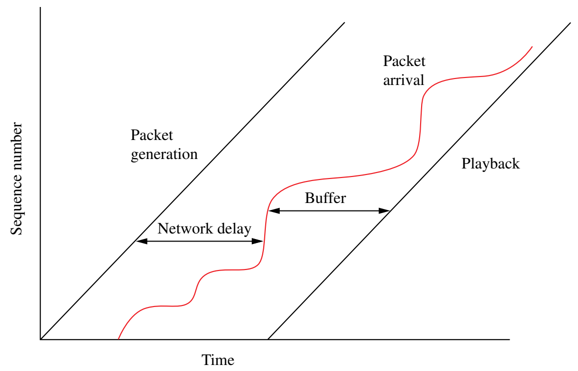
Realtime Applications

- Require "deliver on time" assurances
  - must come from *inside* the network
- Example realtime application (audio)

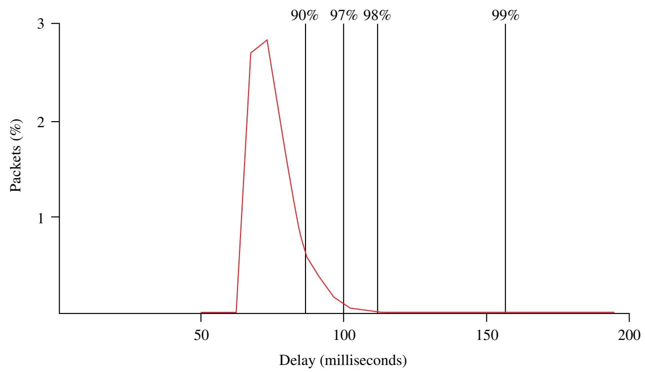


  - sample voice once every 125μs
  - each sample has a *playback time*
  - packets experience variable delay in network
  - add constant factor to playback time: *playback point*

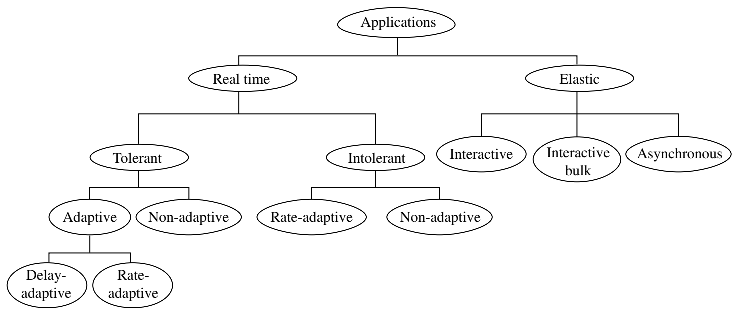
- Playback buffer



- Example distribution of delays



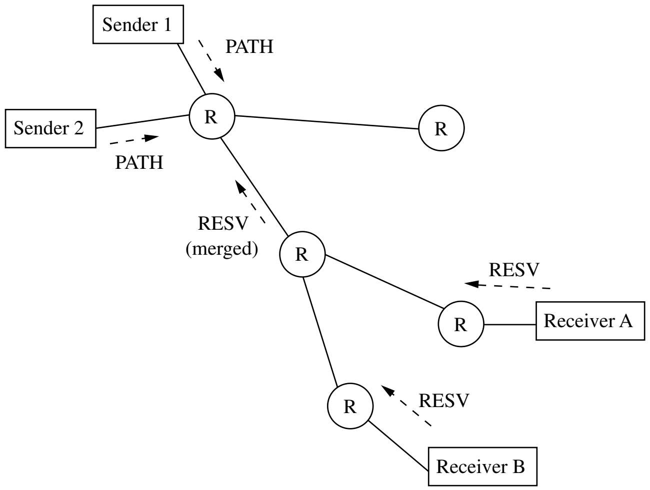
Taxonomy



Integrated Services

- Service Classes
  - guaranteed
  - controlled-load
- Flowspecs
  - *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

- Admission Control
  - decide if a new flow can be supported
  - answer depends on service class
  - not the same as *policing*
- Reservation Protocol
  - called signalling in ATM
  - proposed Internet standard: RSVP
  - consistent with robustness of today's connectionless model
  - uses soft state (refresh periodically)
  - designed to support multicast
  - is receiver oriented
  - two messages: PATH and RESV
  - source transmits PATH messages sent every 30 seconds
  - destination responds with RESV message
  - merge requirements in case of multicast
  - can specify number of speakers



- Packet Processing
  - classification: associate each packet with the appropriate reservation
  - scheduling: manage queues so each packet receives the requested service
- 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: support two classes of packets
  - premium
  - best-effort
- Mechanisms
  - packets: ‘in’ and ‘out’ bit
  - edge routers: tag packets
  - core routers: RIO (RED with In and Out)