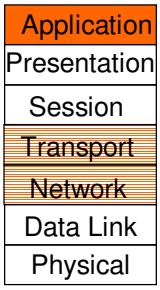
CSE/EE 461 – Lecture 20

Congestion Avoidance

David Wetherall
djw@cs.washington.edu

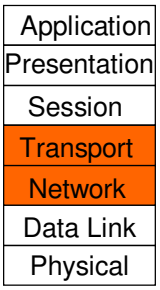
Last Time …

- Introduction to Quality of Service
- Focus
  - What transports do applications need?
- Topics
  - Real-time versus Elastic applications
  - Adapting to variable delay
  - Token buckets as bandwidth descriptors



This Lecture

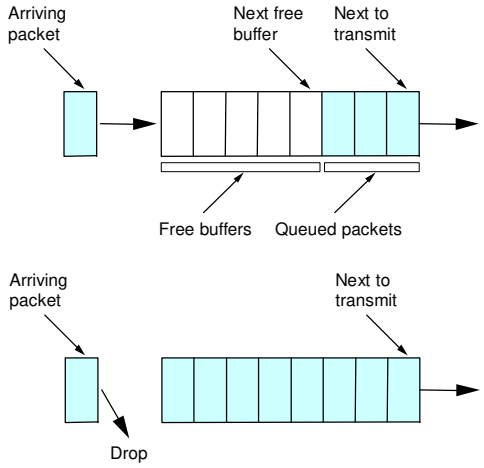
- Congestion Avoidance
- Focus
  - How to we avoid congestion?
- Topics
  - Random Early Detection (RED) gateways
  - Explicit Congestion Notification (ECN)

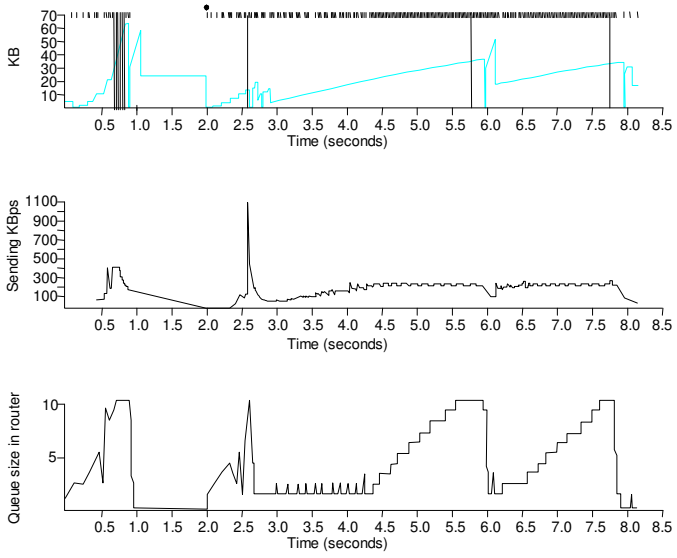


Why Congestion Avoidance?

- TCP causes congestion as it probes for the available bandwidth and then recovers from it after the fact
  - Leads to loss, delay and bandwidth fluctuations (Yuck!)
  - We want congestion avoidance, not congestion control
- Congestion avoidance mechanisms
  - Aim to detect incipient congestion, before loss. So monitor queues to see that they absorb bursts, but not build steadily

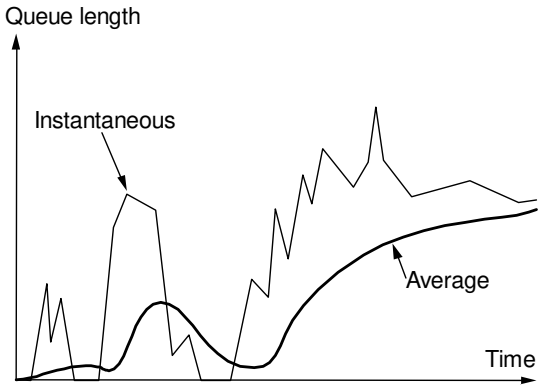
FIFO with Tail Drop





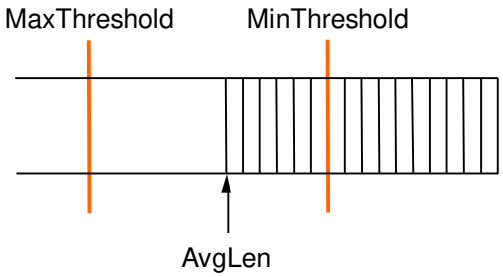
Incipient Congestion at a Router

- Sustained overload causes queue to build and overflow



Random Early Detection (RED)

- Common approach is to have routers monitor average queue and send "early" signal to source when it builds by probabilistically dropping a packet



- Paradox: early loss can improve performance!

Red Drop Curve

- Start dropping a fraction of the traffic as queue builds
  - Expected drops proportional to bandwidth usage
  - When queue is too high, revert to drop tail
  - Nice theory, difficult to set parameters in practice

Explicit Congestion Notification (ECN)

- Why drop packets to signal congestion?
  - Drops are a robust signal, but there are other means …
  - We need to be careful though: no extra packets
- ECN signals congestion with a bit in the IP header
- Receiver returns indication to the sender, who slows
  - Need to signal this reliably or we risk instability
- RED actually works by "marking" packets
  - Mark can be a drop or ECN signal if hosts understand ECN
  - Supports congestion avoidance without loss

Aside: TCP Vegas (Peterson '94)

- RED needs router upgrades but no host upgrades
- Instead, can we upgrade host but not router?

- TCP Vegas looks at the difference between cwnd (the amount of outstanding data in the network) and that acknowledged from the other side in the last interval
  - Excess must be buffered in the network at router queues
  - Vegas slows down when it believes there is a queue and otherwise increases to use the available bandwidth

Key Concepts

- We want to avoid congestion rather than control it after it has occurred
  - Think of in terms of the queues at routers

- Random early packet drops, rather than tail drop, can have unintuitive advantages
  - Signal congestion early, before we're forced to drop repeatedly

- ECN signals congestion using bit in the IP header
  - No loss and no extra packets at overloaded times