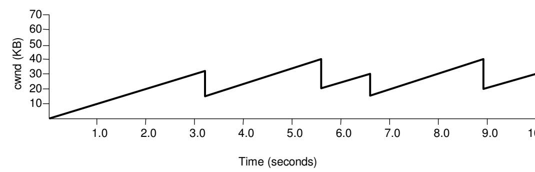
CSE/EE 461 – Lecture 18
Congestion Avoidance

Janet Davis
jlnd@cs.washington.edu
February 18, 2004
Reading: Peterson 6.4

Last time…

- "TCP is self-clocking and the congestion window oscillates around the bottleneck bandwidth."
  - Slow start
  - Additive increase/multiplicative decrease
  - Fast retransmit and fast recovery

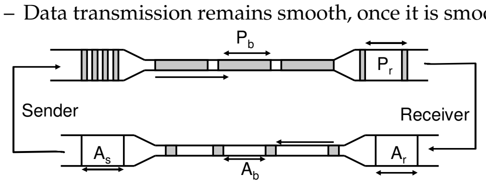


This time…

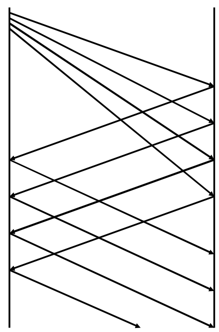
- Loose ends from last time
  - Self-clocking
  - Example (with fast retransmit)
  - Other questions
- Congestion avoidance
  - How can we avoid congestion rather than keeping it under control after it happens?
- Topics:
  - Random Early Detection (RED) queues
  - Explicit Congestion Notification (ECN)

TCP is "Self-Clocking"

- Congestion window helps to "pace" the transmission of data packets
- In steady state, a packet is sent when an ACK is received
  - ACK signals that a packet has left the network and it's safe to send another
  - Data transmission remains smooth, once it is smooth

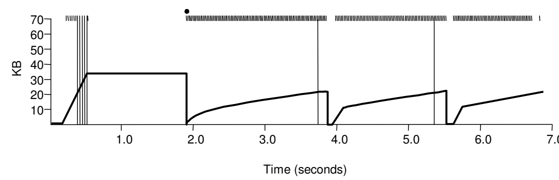


Self-clocking: Time-sequence diagram



L18.5

Example (with Fast Retransmit)



L18.6

Questions from last time…

- Is it normal to have so many losses?
- Are there more efficient ways to deal with congestion?

- Are there priority schemes in use?
- What's the best way to transmit real-time data?

L18.7

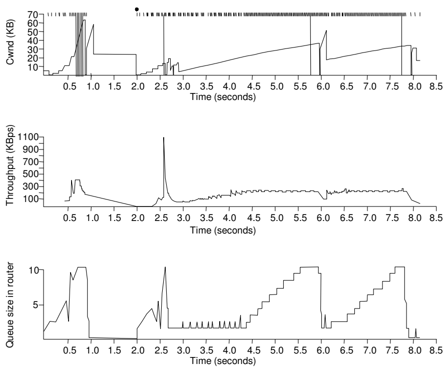
Why Congestion Avoidance?

- TCP probing causes congestion!
  - Leads to loss, delay, and variation in throughput ☹
  - What does this look like?

- We want congestion avoidance, not just congestion control.

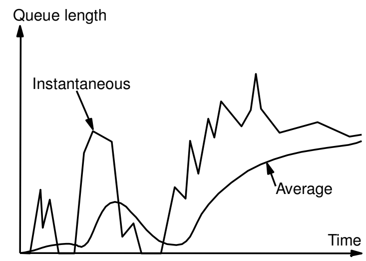
L18.8

TCP Congestion Control with FIFO/Drop-Tail Queues



Incipient Congestion at a Router

- Sustained overload causes average queue length to increase



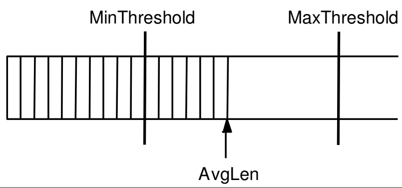
Congestion Avoidance Ideas

- TCP lets us react to congestion after it happens.
- How can we prevent the queue from building up in the first place?
  – Routers? Hosts? Both working together?

Congestion Avoidance Ideas

Random Early Detection (RED)

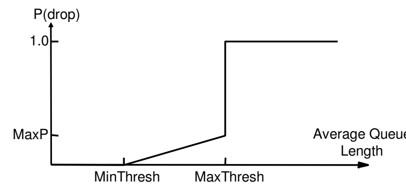
- Idea: Send an "early" signal of congestion when average queue length is high
- AvgLen = Exponentially weighted moving average of queue length
- Probabilistically drop packets when MinThreshold ≤ AvgLen ≤ MaxThreshold
- Paradox: Early loss can improve performance!



L18.13

RED Drop Curve

- Drop a fraction of the traffic as queue builds
  - Expected drops per flow proportional to bandwidth usage
  - When buffer is full, revert to drop tail
  - Nice theory, difficult to set parameters in practice



L18.14

Signalling Congestion

- Why drop packets to signal congestion?
  - Drops are a robust signal, but also a severe one
  - How can we signal congestion without loss and retransmissions?

L18.15

Explicit Congestion Notification (ECN)

- Idea: Use a bit in the IP header to signal congestion
- Receiver reflects signal back to sender
  - Need to signal this reliably or we risk instability
- RED actually works by "marking" packets
  - Mark can be a drop or ECN signal
  - Supports congestion avoidance without loss

L18.16

Congestion avoidance without router support

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

TCP Vegas

- Idea: Compare cwnd/RTT to observed throughput
  - Difference must be buffered in the network at router queues
    - Aim for 1-3 extra packets in network
    - Increase linearly if difference < 1
    - Decrease linearly if difference > 3
  - Still do multiplicative decrease on loss

Active Research Topics

- Active Queue Management
- Equation-based Congestion Control (Peterson 6.5.5)
- Achieving high throughput on high bandwidth-delay product paths
  - To send 10 Gbps with RTT=100ms, can only experience congestion once every 100 minutes!
  - AIMD is just too slow to ramp up
  - eXplicit Control Protocol [Dina Katabi et al. 2002]
  - HighSpeed TCP [Sally Floyd 2003, RFC 3649]

Key Concepts

- We want to avoid congestion rather than control it after it has occurred
- Random early packet drops, rather than tail drop, can have unintuitive advantages
  - Signal congestion early, before we're forced to drop many packets
- ECN signals congestion using bit in the IP header