Random Early Detection (RED) gateways

Sally Floyd
CS 268: Computer Networks
floyd@ee.lbl.gov

March 20, 1995

The Environment

- Feedback-based transport protocols (e.g., TCP)

Problems with current Drop-Tail gateways

- If queue size (buffer allocated for queue) is small, difficulties for large-window TCP connections during slow-start and little ability to accomodate transient congestion. If queue is large, unnecessary delay for delay-sensitive traffic (e.g., telnet TCP).

- Global synchronization, resulting in reduced aggregate throughput [Zhang and Clark, 1990].

- Biases against bursty traffic [Floyd and Jacobson, 1992].

- Fairness concerns, phase effects, etc.

Underlying problems of Drop-Tail gateways

- Transient and persistent congestion are not distinguished.
- Gateway has no control over packet drops (congestion feedback).

Random-Drop gateways give the gateway some control of *which* packet to drop, but do not take care of all of the problems of Drop-Tail gateways.

Approach of RED gateways

- Distinguish between transient and persistent congestion. Design the network to accomodate bursty traffic, rather than shaping bursty traffic to accomodate the needs of the network.

- The most effective detection of persistent congestion can occur at the gateway.

- Let the gateway control packet drops.

- When possible, use FIFO scheduling within a traffic class. FIFO has low overhead, no scaling problems, and reduces the tail of the delay distribution.

- Allow for gradual deployment, flexibility, and evolution.

RED gateway algorithm:

```
for each packet arrival
   calculate the new average queue size avg
   if min_th ≤ avg < max_th
      calculate probability p_a
      with probability p_a:
         mark/drop the arriving packet
   else if max_th < avg
      drop the arriving packet
```

Variables:

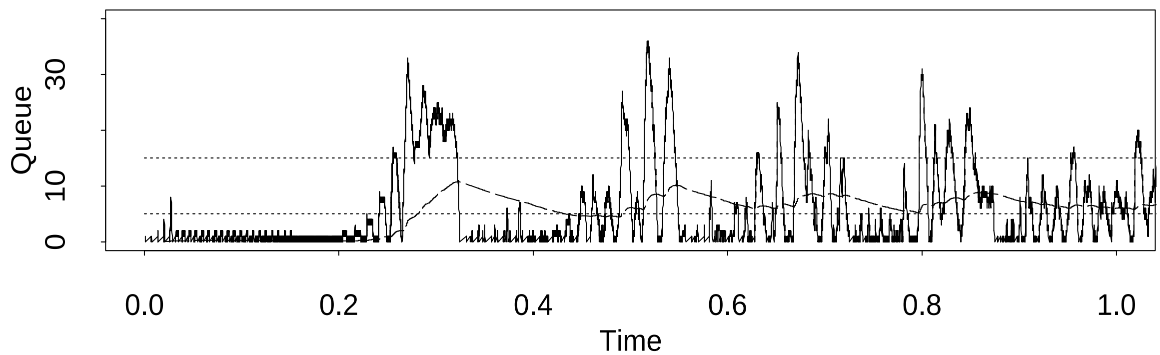
avg: average queue size

p_a: packet marking/dropping probability

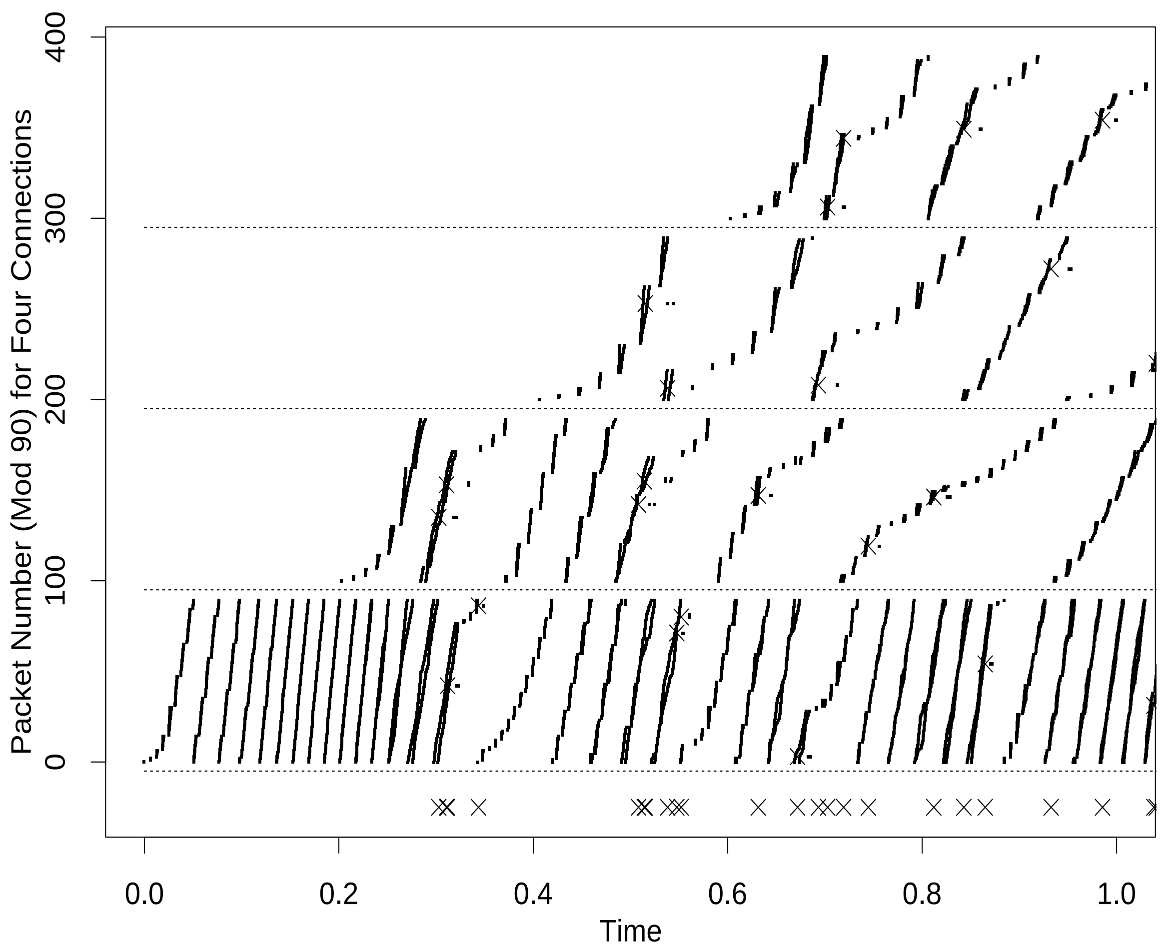
Fixed parameters:

min_{th}: minimum threshold for queue

max_{th}: maximum threshold for queue



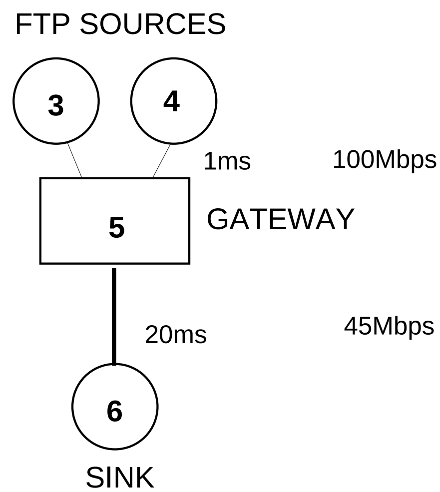
Queue size (solid line) and average queue size (dashed line).

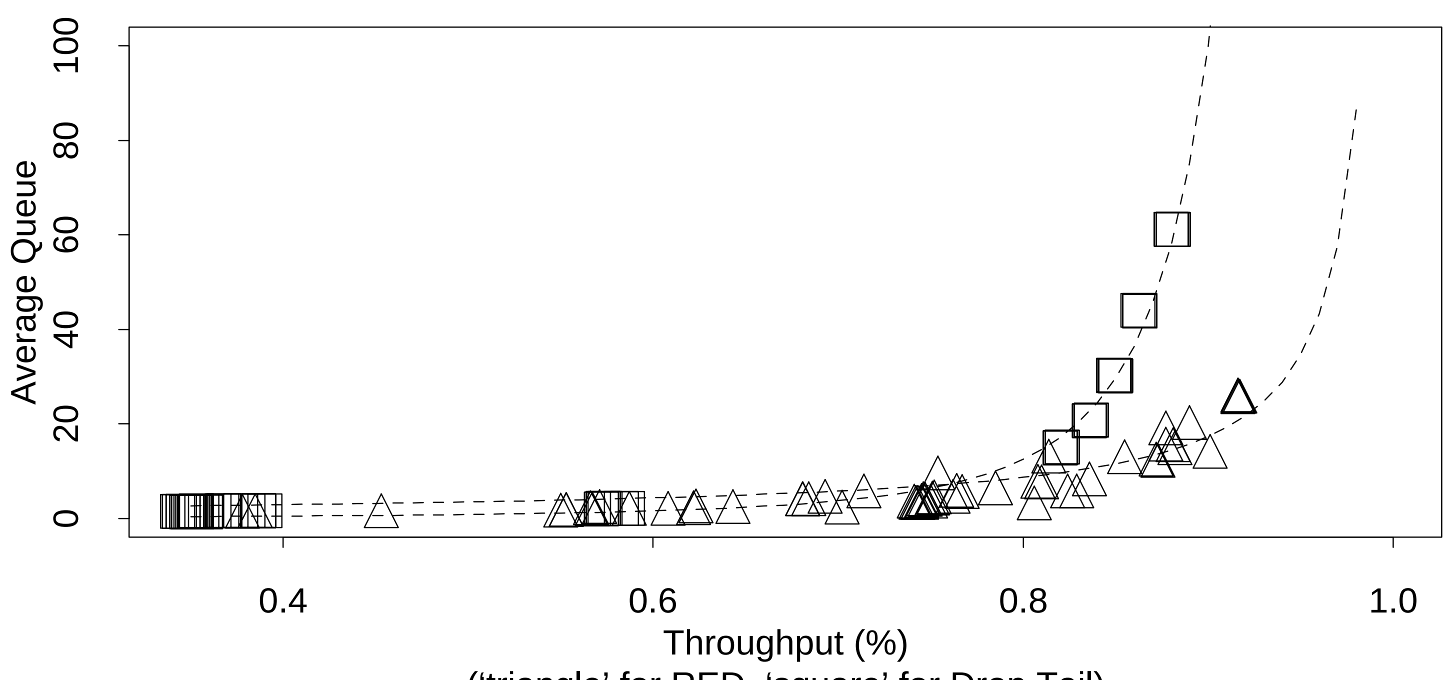


A simulation with four FTP connections with staggered start times.

Comparing Drop Tail and RED gateways.

Goals: high throughput, low average delay.





(‘triangle’ for RED, ‘square’ for Drop Tail)

Calculating the average queue size

```
for each packet arrival
  if the queue is nonempty
```
$$avg \leftarrow (1 - w_q) avg + w_q\, q$$
```
  else using a table lookup:
```
$$avg \leftarrow (1 - w_q)^{(time - q\_time)/s}\, avg$$

Parameters:

- w_q: queue weight
- s: typical transmission time
- q: current queue size
- $time$: current time
- q_time: start of the queue idle time

For w_q a power of two (e.g., $2^{-9} \approx 0.002$), this is efficient to implement.

Time constant for calculating the average queue size

- The averaging has a time constant of $-1/\ln(1 - w_q)$ packet arrivals. That is, if the actual queue size changes from 0 to 1, and then stays at 1, then after $-1/\ln(1 - w_q)$ packet arrivals avg will reach $0.63 = 1 - 1/e$ of the new value.

- To allow a burst of L packets to arrive instantaneously, forming a queue of size L, before the average queue size reaches the minimum threshold min_{th}:

$$L + 1 + \frac{(1 - w_q)^{L+1} - 1}{w_q} < min_{th}.$$

(For min_{th} =5 packets, and L=50 packets, choose $w_q \leq 0.0042$.)

Calculating the packet drop probability

- Method 1: The packet marking probability p is a linear function of avg.

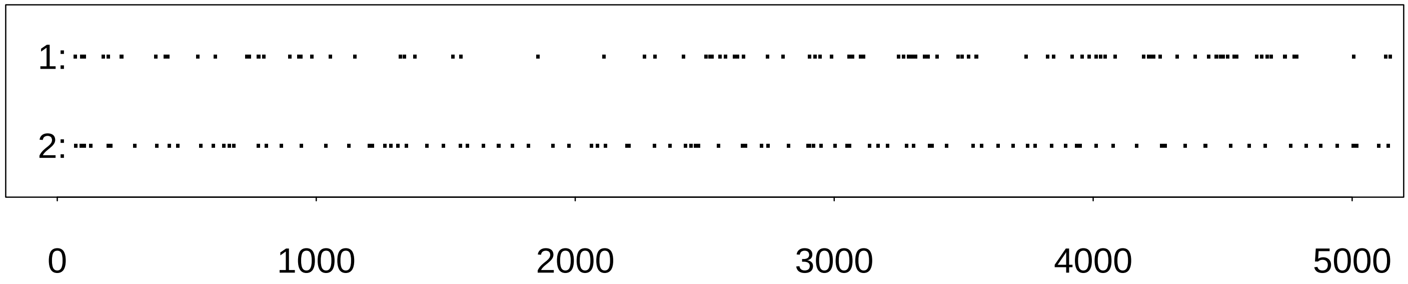
$$p_b \leftarrow max_p(avg - min_{th})/(max_{th} - min_{th})$$

This gives geometric (roughly exponential) intermarking intervals (the number of packets between marked/dropped packets).

- Method 2: The packet marking probability increases as $count$, the number of packets since the last marked packet, increases.

$$p_a = p_b/(1 - count \cdot p_b)$$

This gives intermarking intervals uniform on $[1, 1/p_b]$.



(top row for Method 1, bottom row for Method 2)

Misbehaving users

- The probability that the gateway chooses a particular connection to notify (e.g., mark or drop) during congestion is roughly proportional to that connection's share of the bandwidth at the gateway.

- If a connection has a large fraction of the recent marked packets, then it has most likely received a large fraction of the recent bandwidth.

- This gives an easy method for identifying high-bandwidth users in times of congestion. If desired, such high-bandwidth users could receive differential treatment.

Implementations

RED algorithms implemented in software for IBM RS/6000 routers on ANSNET (45 Mbps network between Hawaii and continental US). The main goal is to achieve high link utilization, including for high-bandwidth TCP connections to the Maui High Performance Computing Center.

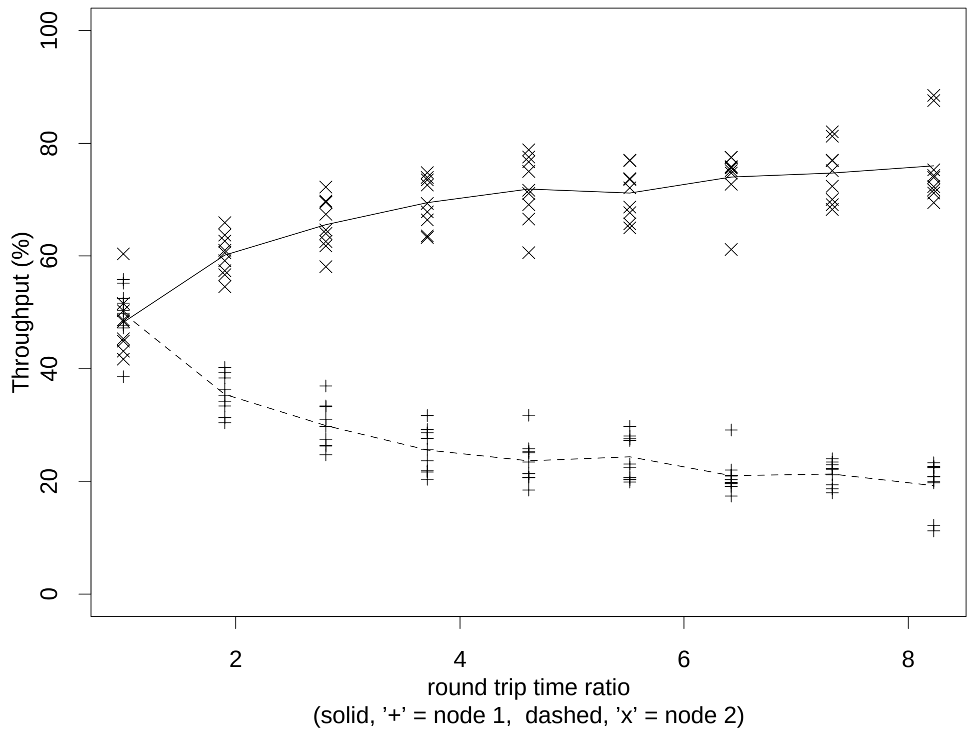
Open questions

- The optimal average queue size?
- RED gateways with ATM networks?
- RED gateways with rate-adaptive video?
- More on misbehaving users.

Fairness issues with RED gateways:

- No bias against bursty traffic.

- TCP traffic has a bias against connections with longer RTTs, due to the current TCP window increase algorithm [Floyd and Jacobson, 1992].



- There is a bias against connections that pass through multiple congested gateways [Floyd 1991]. It is not clear (to me) that this is a bad thing.

- Treatment of misbehaving users?

TCP and Explicit Congestion Notification?

- Advantages of ECN over packet drops:
  - Avoiding unnecessary packet drops.
  - Quicker detection by sources of congestion (without having to wait to detect a dropped packet).

- Disadvantages of ECN:
  - Lost ECN messages (e.g., Source Quench, lost ACK packets with congestion indication bits)
  - Non-compliant sources?

- Changes to TCP: Response should be similar, over longer time scales, to the response to a dropped packet. A single ECN would be treated as an indication of congestion, but TCP would react at most once per RTT [Floyd 1994].

References

- (Zhang and Clark 1990), Zhang, L., and Clark, D., "Oscillating Behavior of Network Traffic: A Case Study Simulation", *Internetworking: Research and Experience*, Vol. 1 No. 2, December 1990, pp. 101-112.
- (Floyd 1991) *Connections with Multiple Congested Gateways in Packet-Switched Networks Part 1: One-way Traffic*, Computer Communication Review, V.21 N.5, October 1991, pp. 30-47.
- (Floyd and Jacobson 1992), *On Traffic Phase Effects in Packet-Switched Gateways*, Internetworking: Research and Experience, V.3 N.3, September 1992, p.115-156.
- (Floyd and Jacobson 1993), *Random Early Detection Gateways for Congestion Avoidance*, IEEE/ACM Transactions on Networking, V.1 N.4, August 1993, p. 397-413.
- (Floyd 1994), "TCP and Explicit Congestion Notification", to appear in ACM Computer Communication Review.