CS644
Advanced Networks

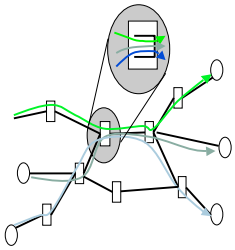
Lecture 5

AQM

Andreas Terzis

Router Support For Congestion Management

- Traditional Internet
  - Congestion control mechanisms at end-systems, mainly implemented in TCP
  - Routers play little role
- Router mechanisms affecting congestion management
  - Scheduling
  - Buffer management
- Traditional routers
  - FIFO
  - Tail drop



Drawbacks of FIFO with Tail-drop

- How to pick buffer size?
  - Small router buffers: TCP connections get many losses during slow start
  - Large buffers: unnecessary delay
- Buffer lock out by misbehaving flows
- Synchronizing effect for multiple TCP flows [ZSC91]
- Burst or multiple consecutive packet drops
  - Bad for TCP fast recovery

Underlying problem with Drop Tail

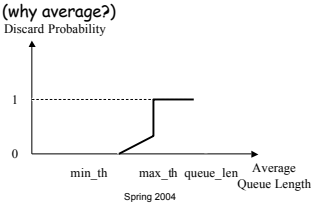
- No way to differentiate between transient and persistent congestion
- Router has no control over packet drops

RED Approach

- Distinguish between transient and persistent congestion
  - Design the network to accommodate bursty traffic
  - Gateway is the most effective decision point for persistent congestion
  - Use FIFO scheduling
    - Low overhead, good scaling characteristics, reduces delay
  - Allow gradual deployment

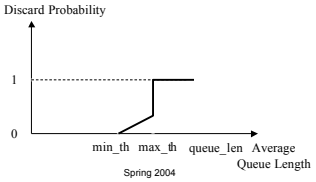
RED

- FIFO scheduling
- Buffer management:
  - Probabilistically discard packets
  - Probability is computed as a function of average queue length (why average?)



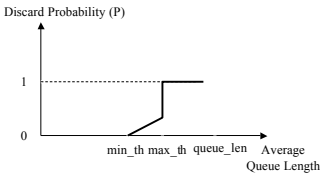
RED (cont'd)

- min_th – minimum threshold
- max_th – maximum threshold
- avg_len – average queue length
  - avg_len = (1-w)*avg_len + w*sample_len



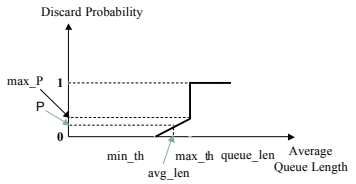
RED (cont'd)

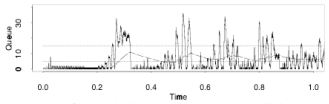
- If (avg_len < min_th) → enqueue packet
- If (avg_len > max_th) → drop packet
- If (avg_len >= min_th and avg_len < max_th) → enqueue packet with probability P



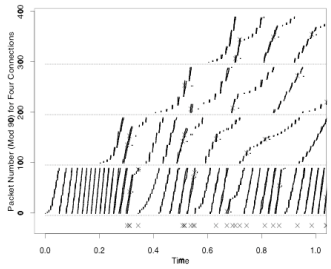
RED (cont'd)

- P = max_P*(avg_len – min_th)/(max_th – min_th)
- Improvements to spread the drops
  - P' = P/(1 – count*P), where
    - count – how many packets were consecutively enqueued since last drop

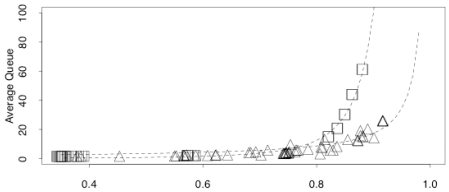




Queue size (solid line) and average queue size (dashed line).



Comparison with Drop Tail



('triangle' for RED, 'square' for Drop Tail)

Misbehaving Users

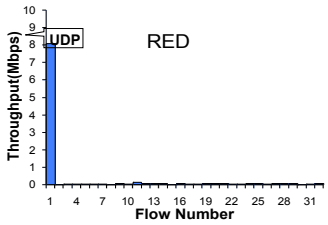
- The probability that a connection is notified is proportional to the connection's share of the router's bandwidth
- Packet drops can be used to identify misbehaving users
  - Can be used to further penalize them
    - Penalty Box
    - Pushback

RED Advantages

- Absorb burst better
- Avoids synchronization
- Signal end systems earlier

Problems with RED

- No protection: if a flow misbehaves it will hurt the other flows
- Example: 1 UDP (10 Mbps) and 31 TCP's sharing a 10 Mbps link



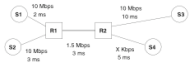
Promoting E2E Congestion Control [FF99]

- Congestion control was critical factor to the success of the Internet
  - Network no longer a small testbed for friendly researchers
    - Fundamental change that has multiple consequences
    - Network has to take *active role* in protecting from misbehaving end users
  - Possible solutions
    - Deploy *isolation* mechanisms (e.g. Fair Queuing)
    - Provide incentives for continued use of e2e congestion control
    - Pricing mechanisms
      - Can mix and match

What does e2e congestion control provide?

- Avoid congestion collapse
- Provide Fairness

Unfairness example



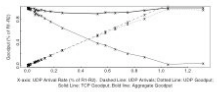
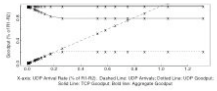


Figure 2: Simulations showing extreme unfairness with three TCP flows and one UDP flow, and FCFS scheduling.



About congestion collapse

- Classical Congestion Collapse
  - Paths clogged with unnecessarily-retransmitted packets
  - Fix: TCP retransmit timers and congestion control algorithm
- Congestion collapse from undelivered packets
  - Paths are clogged with packets that are discarded before they reach the receiver
  - Due to applications not using E2E congestion control

Goodput

- Bandwidth delivered to receiver excluding duplicate packets

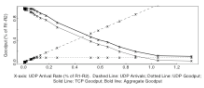
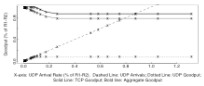


Figure 4: Simulations showing congestion collapse with three TCP flows and one UDP flow, with FCFS scheduling.



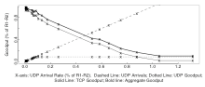
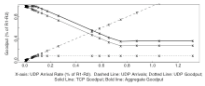
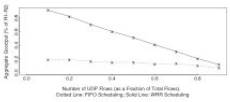


Figure 6: Simulations with one TCP flow and three UDP flows showing congestion collapse with FIFO scheduling.



Limits of scheduling discipline

- WRR/FQ limit the resources individual flows receive
  - Not helpful when large percentage of flows are misbehaving



Building the right incentives

- What do we need in the network architecture to induce applications to employ e2e congestion control?
- Discover unresponsive flows and regulate them
  - How to discover these flows
  - How to regulate them