Router Support For Congestion Management

- Traditional Internet
  - Congestion control mechanisms at end-systems, mostly 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

\[ \text{avg}_\text{len} = (1-w) \times \text{avg}_\text{len} + w \times \text{sample}_\text{len} \]
RED (cont’d)

- \( P = \max_P \times \left( \frac{\text{avg}_\text{len} - \text{min}_\text{th}}{\text{max}_\text{th} - \text{min}_\text{th}} \right) \)

- Improvements to spread the drops
  \( P' = P/(1 - \text{count}^P) \), where
  \( \text{count} \) - how many packets were consecutively enqueued since last drop.

<table>
<thead>
<tr>
<th>Discrim Probability</th>
<th>Average Queue Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \max_P )</td>
<td>( \text{avg}_\text{len} )</td>
</tr>
<tr>
<td>( \min_\text{th} )</td>
<td>( \text{max}_\text{th} )</td>
</tr>
</tbody>
</table>

Comparison with Drop Tail

- 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

Misbehaving Users
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

![Graph showing throughput vs. flow number for UDP and RED]

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

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

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