Experimental Evaluation of the Effect of Queue Management Schemes on the Performance of High Speed TCPs in 10Gbps Network Environment

Lin Xue*, Cheng Cui, Suman Kumar, Seung-Jong Park

Computer Science Department and Center for Computation and Technology Louisiana State University
Outline

- Background
- Motivation
- Experimental Design
- Experimental Results and Discussion
- Conclusion and Future Research Direction
Problem environment:

- Internet 2, National Lambda Rail (NLR), LONI, etc.
- Few number of users and several parallel TCP flows
- Smaller size buffer than BDP
- The classic TCP “sawtooth”
- High burst traffic
Previous Works

- Individual evaluation over 10Gbps *
  - Factor 1: Queue management schemes
    - Drop-tail, RED, CHOKe, SFB, etc.
  - Factor 2: Queue size
    - 1% to 100% of BDP
  - Factor 3: Congestion avoidance schemes
    - TCP-Reno, BIC/CUBIC, HSTCP, etc.

- Inter-relationship among factors should be evaluated

Motivation & Goal

- Network operator’s Dilemma
  - Which queue management schemes to use
  - How much buffering to provide

- Network Users Dilemma
  - Which high speed TCP variants to use

Goal:
- Understand the impact of queue management schemes and router buffer size on the performance of high speed TCPs
- The effect of these three on the performance of 10Gbps high speed networks
Testbed Setup

- Network topology = Dumbbell
- Number of flows = 10 from each sender
- Link bandwidth = all links have 10Gbps
- Link delay = 120ms

- TCPs = TCP-Reno/CUBIC/HSTCP
- Queue = Drop-tail/RED/CHOKe/SFB
- Queue sizes = 1%/5%/10%/20%/40%/60%/100% of BDP

Fig. 1. Experiment Topology
Evaluation Methods

- Three main methods for evaluation.
  - Simulations
    - Network Simulator 2 (NS-2)
  - Experiments over emulated networks
    - Emulab (100Mbps)
  - Experiments over production networks
    - PlanetLab (100M/1Gbps), etc.

- None of them scale to bandwidth of the order of 10Gbps.

- We developed a systematic and repeatable evaluation method in a 10Gbps high speed network environment:
  - CRON: Cyber-infrastructure of Reconfigurable Optical Network
  - http://www.cron.loni.org
The link utilization is the percentage of total bottleneck capacity (10Gbps) utilized during an experiment run.

- 1% BDP buffer size -> more than 85% link utilization
- 10% BDP buffer size -> more than 90% link utilization
- Drop-tail -> good for TCP-Reno
- SFB -> good for HSTCP

Fig. 2. Link Utilization as a function of buffer size in each TCP variant
Intra-Protocol Fairness

We fix buffer size of the bottleneck router to 10% BDP, and calculate Jain’s fairness index among 20 TCP flows.

- RED shows highest fairness
- Drop-tail performs bad, especially with TCP-Reno
- Active queue management schemes get more fairness than drop-tail

<table>
<thead>
<tr>
<th></th>
<th>Drop-tail</th>
<th>RED</th>
<th>CHOKe</th>
<th>SFB</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUBIC</td>
<td>0.9884</td>
<td>0.9943</td>
<td>0.9807</td>
<td>0.9813</td>
</tr>
<tr>
<td>HSTCP</td>
<td>0.9782</td>
<td>0.9872</td>
<td>0.9897</td>
<td>0.9729</td>
</tr>
<tr>
<td>RENO+SACK</td>
<td>0.9356</td>
<td>0.9772</td>
<td>0.9704</td>
<td>0.9758</td>
</tr>
</tbody>
</table>

**TABLE I**
Jain’s Fairness Index among TCP flows
RTT Fairness

- Fairness between two flows with different RTTs
  - Sender1: RTT = 120ms
  - Sender2: RTT varies as 30, 60, 120, and 240ms.

- CUBIC has the highest RTT fairness than other TCPs
- SFB provides good RTT fairness in general

Fig. 3. RTT fairness as a function of RTT in each TCP variant
Delay

- Delay = propagation delay (120ms) + queuing delay
  - In CUBIC, Drop-tail has the longest delay
  - In HSTCP and RENO, SFB has more delay
  - RED and CHOKe keep low delay

Fig. 4. Delay as a function of buffer size in each TCP variant
Packet Drop Rate

- Packet drops include:
  - Drops at middle of queue because of active operation of AQMs
  - Drops at tail of queue because of buffer overflow.

- SFB’s drop rate is high with CUBIC
- RED’s drop rate is low with CUBIC or Reno
- Drop-tail’s drop rate is low with HSTCP, but high with Reno

![Graphs showing Packet Drop Rate vs. buffer size for CUBIC, HSTCP, and RENO](image-url)
Complexity

- Average memory consumption on the bottleneck router
  - SFB induces high memory consumption, especially with HSTCP
  - RED keeps memory consumption low
  - Drop-tail has more memory consumption than RED and CHOKe, but less than SFB

<table>
<thead>
<tr>
<th>Queue Scheme</th>
<th>Drop-tail</th>
<th>RED</th>
<th>CHOKe</th>
<th>SFB</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUBIC</td>
<td>224</td>
<td>45</td>
<td>51</td>
<td>420</td>
</tr>
<tr>
<td>HSTCP</td>
<td>197</td>
<td>39</td>
<td>37</td>
<td>1015</td>
</tr>
<tr>
<td>RENO+SACK</td>
<td>202</td>
<td>91</td>
<td>108</td>
<td>167</td>
</tr>
</tbody>
</table>

**TABLE II**

_Average Memory consumption on Router (MB)_

- The CPU usage results reveal less than 10% of total CPU usage for all queue schemes
Conclusion and Future Work

- A preliminary experimental evaluation on the effect of inter-relationship among:
  - Queue management schemes (Drop-tail, RED, CHOKe, and SFB)
  - Queue size
  - High speed TCPs (TCP-Reno, CUBIC and HSTCP)

- Future work
  - To observe the impacts on a more realistic experimental environment with background traffic.
  - To study the behavior of heterogeneous TCP flows mixed in the bottleneck.
Questions ?
Experiment with CRON

- Experimental design with Java based GUI of Emulab
  - Additional features such as tracing, Link Queuing policy, traffic generators, availability of TAR files etc.
Experiment with CRON contd…
Experiment with CRON contd…

- Y-topology similar to Dumbbell
- Dummynet software emulators were used to emulate large size buffers
- Bottleneck link has 8Gbps bandwidth and 30msec