The War Between
Mice & Elephants

by,
Matt Hartling
&
Sumit Kumbhar

ACN Presentation
Feb. 5, 2002

Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work
Short vs. Long Lived Flows

**Long-Lived (Elephants)**
- Transmit large number of packets
- Operate primarily in TCP congestion avoidance
- TCP mechanisms designed for long-lived flows

**Example:**
- FTP

**Short-Lived (Mice)**
- Transmit small number of packets
- Operate Primarily in slow-start phase

**Example:**
- HTTP/1.0 transfers
HTTP/1.0
Web Traffic (HTTP/1.0)

Bruce Mah’s Statistics (HTTP/1.0; 1995)

- HTTP Reply Lengths:
  - Max \(=\) 1 MB
  - Mean \(=\) 8-10 KB
  - Median \(=\) 2 KB (50% send only 1 or 2 packets!!)

This paper uses BU Traces from 1995 (old data!)

HTTP/1.0 reply lengths modeled using Pareto distribution.
Short-Lived Flow Issues (1): Slowstart

- Transmission rate increases slowly
- Added delay
- Bad for flows that don’t get out of the early stages of slowstart.
Short-Lived Flow Issues (2): Small cwnd

- Fast Retransmit needs three duplicate acks
- If one of the first three packets are dropped, a TO occurs
- TO’s are BAD!! (Web traffic is somewhat more sensitive to delays.)
Short-Lived Flow Issues (3): Initial Timeout Value

- Initial timeout value (ITO) is set to 3 seconds
- If the SYN, SYN-ACK, or first data packet is dropped, the flow must wait for the 3 second TO before a retransmission
Related Work

Class Based Isolation of short, long, and UDP [Yilmaz, Selma; Matta, Ibrahim 2001]

TCP Protocol Solutions:
- Reduce ITO
- Larger Initial CWND

“Size-aware job scheduling” [Crovella et al. 1999 and Bansal et al. 2001]
Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work
Average Transmission Time

(a) Average Transmission Time
Transmission Time Variance

(b) Coefficient of Variation
Comparison of Drop Tail, RED, RIO-PS

Fig. 2. Impact of Preferential Treatment—Link utilization under Drop Tail (left), RED (middle), and RIO-PS (right)
## Goodput

<table>
<thead>
<tr>
<th>Link B/W</th>
<th>Flows</th>
<th>DropTail</th>
<th>RED</th>
<th>RIO-PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.25Mbps</td>
<td>All</td>
<td>153479</td>
<td>154269</td>
<td>154486</td>
</tr>
<tr>
<td></td>
<td>Short</td>
<td>40973</td>
<td>49897</td>
<td>49945</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>112506</td>
<td>104372</td>
<td>104541</td>
</tr>
<tr>
<td>1.5Mbps</td>
<td>All</td>
<td>185650</td>
<td>184315</td>
<td>183154</td>
</tr>
<tr>
<td></td>
<td>Short</td>
<td>43854</td>
<td>49990</td>
<td>49990</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>141796</td>
<td>134325</td>
<td>133164</td>
</tr>
</tbody>
</table>

**TABLE I**

**NETWORK GOODPUT UNDER DIFFERENT SCHEMES**
Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work
RIO-PS

RIO-SL Router (core)
- Uses RIO RED
- Less Drops Short-Lived Flows
- Uses separate RED parameters for each Class ($\text{min}_I$, $\text{max}_I$, and $\text{max}_p$)

Edge Router
- Classifies Packets as Short or Long
- Perflow State Packet Count
Edge Router Functions

- Maintains per flow packet counts
- Labels the packet as “short” or “long” based on the following parameters:
  - $L_t$ - threshold when exceeded packets are labeled as long (Dynamic or Static)
  - $T_u$ – idle timer; if no packets are received within time $T_u$ seconds, the flow is removed from the system
Edge Router Functions (cont)

Dynamic \( L_t \) parameters:

- \( SLR \) – target ratio \# short flows / \# long flows.
- \( T_c \) – time between making additive adjustments to \( L_t \) to achieve SLR

All Long flows begin as Short
Core Router Functions

- Uses RIO Mechanism
- RIO (David Clark – MIT and Wenjia Fang – Princeton University)
  - Sender or receiver classifies the packet as “in” or “out” of profile
    - (Profile example: transmission rate of 64K)
  - Basic idea – drop out of profile packets more aggressively than in profile.
  - Maintains 2 sets of RED parameters (minth, maxth, maxp, wq, etc.)
Core Router Functions (cont)

Fig. 4. RIO queue with Preferential treatment to Short fbws
Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work
RIO with preference to short flows

Exp 1: \( x = 8 \)
Exp 2: \( x = y = 4.5 \)

15ms
100Mb

all access links
[0.1 - 1] ms
10 Mbps
<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packet Size</td>
<td>500 bytes</td>
</tr>
<tr>
<td>Maximum Window</td>
<td>128 packets</td>
</tr>
<tr>
<td>TCP version</td>
<td>Newreno</td>
</tr>
<tr>
<td>TCP timeout Granularity</td>
<td>0.1 seconds</td>
</tr>
<tr>
<td>Initial Retransmission Timer</td>
<td>3.0 seconds</td>
</tr>
<tr>
<td>B/W delay product (BDP)</td>
<td>≈ 200 pkts (Exp1)</td>
</tr>
<tr>
<td></td>
<td>≈ 120 pkts (Exp2)</td>
</tr>
<tr>
<td>Bottleneck</td>
<td>DropTail: 1.5× BDP</td>
</tr>
<tr>
<td>Buffer Size (B)</td>
<td>RED/RIO-PS: 2.5×BDP</td>
</tr>
<tr>
<td><strong>Q. Parameters</strong></td>
<td></td>
</tr>
<tr>
<td>RED</td>
<td>(0.15B, 0.5B, 1/10, 1/512)</td>
</tr>
<tr>
<td>RIO-PS short</td>
<td>(0.15B, 0.35B, 1/20, 1/512)</td>
</tr>
<tr>
<td>RIO-PS long</td>
<td>(0.15B, 0.5B, 1/10, 1/512)</td>
</tr>
<tr>
<td>RED &amp; RIO-PS</td>
<td>ecn_ on, wait_ on, gentle_ on</td>
</tr>
<tr>
<td>Edge Router</td>
<td>$SLR = 3, T_i = 1$ sec, $T_o = 10$ sec</td>
</tr>
<tr>
<td><strong>Foreground Traffic</strong></td>
<td></td>
</tr>
<tr>
<td>(Src, Dest)</td>
<td>(Server Pool, Client Pool)</td>
</tr>
<tr>
<td>Long Connection Size</td>
<td>1000 packets</td>
</tr>
<tr>
<td>Short Connection Size</td>
<td>10 packets</td>
</tr>
</tbody>
</table>

**TABLE III**
**NETWORK CONFIGURATION**
Average response time for different sized objects
Initial retransmission timer = 3 seconds
Average response time for different sized objects

Initial retransmission timer = 1 second
Instantaneous Queue Size
Instantaneous Drop/Mark Rate
Fairness Index for Short Connections
Fairness Index For Long Connections

- DropTail
- RED
- RIO-PS
Transmission Time For Short Connections
Transmission Time For Long Connections
### Network Goodput Over The Last 2000 Seconds

<table>
<thead>
<tr>
<th>Scheme</th>
<th>DropTail</th>
<th>RED</th>
<th>RIO-PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp1 (ITO=3sec)</td>
<td>4207841</td>
<td>4264890</td>
<td>4255711</td>
</tr>
<tr>
<td>Exp1 (ITO=1sec)</td>
<td>4234309</td>
<td>4254291</td>
<td>4244158</td>
</tr>
<tr>
<td>Exp2 (ITO=3sec)</td>
<td>4718311</td>
<td>4730029</td>
<td>4723774</td>
</tr>
</tbody>
</table>
Fairness Index of Short Connections
Fairness Index For Long Connections
Transmission Time For Short Connections
Transmission Time For Long Connections
Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work
Discussion

1. Simulation Model
   - All TCP connections have similar end to end propagation delays.
   - Better performance obtained in the presence of reverse traffic.

2. The Queue Management Policy
   - RIO does not provide class based or flow guarantees.
   - Other option could be PI controlled RED queue.
3. **Deployment Issues**
   - Success of the scheme depends on how well the edge router can classify the traffic to be long or short.

4. **Flow Classification**
   - Long connections initially are classified as Short and so are treated with higher preference.
Discussion

5. Controller Design
   - “SLR” depends on
     
     \[ T_c = \text{time after which classification threshold } L_t \text{ is updated.} \]
     
     \[ T_u = \text{time after which active flow table is updated.} \]

6. Malicious Users
   - Breaking long transmissions into short.
   - Problem of overhead.
Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work
Conclusions

- Implementation of TCP service that classifies traffic based on their Duration.
- Performance of Short TCP connections is Improved.
- Performance of first few Long connections is also improved.
- Proposed Architecture is good in the terms of deployment.