A Preliminary Analysis of TCP Performance in an Enterprise Network

INM/WREN'10

Boris Nechaev¹, Mark Allman², Vern Paxson²,³, Andrei Gurtov¹

¹Helsinki Institute for Information Technology/Aalto University
²International Computer Science Institute
³University of California, Berkeley

27 April 2010
Background

- Enterprise traffic remains mostly unexplored
  - Logistically difficult to monitor
  - Enterprises are often viewed as working “well enough”
- Data:
  - Lawrence Berkeley National Laboratory
  - October 2005 – March 2006
  - Captured at switches, often switched to new set of ports
  - 351 distinct hosts monitored (≈ 4% of total)
  - 292 million intra-enterprise TCP packets
  - Non-trivial calibration challenges (IMC'09 paper)
- Focus only on intra-enterprise traffic
- Used Bro 1.5.1 to reconstruct connection status
SF + RSTO + RSTR are “good” connections

- 363K “good” connections (68%)
- 50 GB of data transferred
- Consider only these connections in further analysis
- High percentage variability across traces

REJ connections

- Almost all originate at the same host
- Scanning traffic

OTH connections

- Bro observed neither establishment nor teardown
- Over 90% contain a single ACK or data packet
44% of connections stay inside the subnet

Prevalent applications

- Proportions of bytes/connections are unbalanced
- Dantz backup: 27% bytes, 0.3% connections
- HTTP: 9% bytes, 18% connections
- NetBIOS-SSN: 1.5% bytes, 10% connections

An application may show heavy tail in connection size or not

<table>
<thead>
<tr>
<th>App.</th>
<th>Med.</th>
<th>99th</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dantz</td>
<td>6.4 KB</td>
<td>233 MB</td>
<td>4 GB</td>
</tr>
<tr>
<td>ssh</td>
<td>5.5 KB</td>
<td>19 MB</td>
<td>2.6 GB</td>
</tr>
<tr>
<td>NFS</td>
<td>72 B</td>
<td>1.0 MB</td>
<td>1.1 GB</td>
</tr>
<tr>
<td>HTTP</td>
<td>1.9 KB</td>
<td>82 KB</td>
<td>835 MB</td>
</tr>
<tr>
<td>NetBIOS-SSN</td>
<td>2.0 KB</td>
<td>59 KB</td>
<td>137 MB</td>
</tr>
<tr>
<td>Warewulf</td>
<td>6.6 KB</td>
<td>52 KB</td>
<td>52 KB</td>
</tr>
<tr>
<td>Portmap</td>
<td>92 B</td>
<td>716 B</td>
<td>1.1 KB</td>
</tr>
</tbody>
</table>
Prevalence in terms of pkts

Protocol or application

HTTP
SSH
SNMP
NetBIOS
Dantz Backup
LLC–CGMP
NFS/UDP
LLC–APPLETALK
797/udp
NFS/TCP
ARP
9406/tcp
9409/tcp
9407/tcp
LLC–IPX

Trace

(a) Subnet
Prevalence in terms of pkts

Trace

(b) LBL
• Distribution of connection sizes (bytes)
• Ratio of originator data bytes to responder data bytes

- Median transfer size ~2KB
- 90% of traffic comes from just 160 connections (out of 363)
Performance

- Very low number of packets with bad TCP checksum - 583
- 0.1% connections had packet reordering
- No replicated packets
- 0.5% connections experienced retransmissions
  - (Haven’t done fully robust retransmission detection yet)
- Connection maximum flight sizes
  - Median: 214 bytes
  - 99th percentile: 5.3 KB
  - Bandwidth-delay product for 100Mb/s, 1 msec RTT: 12.5
    - Do we see bandwidth underutilization?
Transfer rates

- Rate = \( \frac{\text{Total bytes in the connection}}{\text{Duration}} \)

- Higher rates compared to WAN traffic studies
- Intra-subnet rates are 10 times higher than inter-subnet
Transfer rates (cont'd)

- 4 types of flows with 10 KB and 1 sec thresholds

<table>
<thead>
<tr>
<th>Type</th>
<th>Conns. (%)</th>
<th>Bytes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-Small</td>
<td>57.2</td>
<td>0.6</td>
</tr>
<tr>
<td>Short-Large</td>
<td>2.6</td>
<td>0.8</td>
</tr>
<tr>
<td>Long-Small</td>
<td>31.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Long-Large</td>
<td>8.4</td>
<td>97.8</td>
</tr>
</tbody>
</table>

Why these don’t go faster merits investigation!

Warewulf (9873/tcp)

NetBIOS (139/tcp)

Repetitive HTTP xfers
Summary

- Preliminary analysis of TCP performance
- Higher rates than in WAN
- Less loss than in WAN
- In general, enterprise connections appear to work well
  - Are flaws masked by high capacity and low delays?
- Next steps:
  - Analysis of packet latency dynamics
  - Assessment of loss & retransmission behavior
  - In-depth study of bandwidth utilization
  - Incorporation of a large new dataset
    - 1,000 end systems recorded 2009/2010