A Proposed Framework for Calibration of Available Bandwidth Measurement Tools

Joel Sommers / University of Wisconsin-Madison / jsomers@cs.wisc.edu
Paul Barford / University of Wisconsin-Madison / pb@cs.wisc.edu
Walter Willinger / AT&T Labs-Research / walter@research.att.com

http://wail.cs.wisc.edu
Introduction

• Many tools have been developed to measure Internet properties of interest
  • packet loss, delay, bottleneck capacity, available bandwidth, router connectivity, AS connectivity

• Examining the validity and accuracy of these tools remains a challenging problem
  • typical approach: ns2 simulations followed by small-scale in situ experiments
Measurement Tool Calibration

- Two key aspects of calibration
  - comparison with a known standard
  - adjustment to match a known standard
- Goals of calibration
  - detect inaccuracy in underlying data
  - detect misconceptions or errors in analysis
  - expose limitations of measurement tools
Focus area: available bandwidth

- Measurement of available bandwidth (AB) has become an area of interest recently
  - Capacity planning, overlay management
- Defined as the minimum spare capacity along a path over some time interval
- Calibrating AB estimation tools
  - Comparison with “ground truth” of sufficient accuracy
  - Adjustment of algorithm and/or parameters
Background: Spruce
(Strauss & Katabi, IMC 2003)

• Assume link of least capacity (narrow link) is the same as the congested (tight) link and that capacity $C$ is known

• Expansion of packet pair is proportional to volume of cross traffic

$$A = C \left( 1 - \frac{g_{out} - g_{in}}{g_{in}} \right)$$

• Send 100 packet pairs according to Poisson process, report average $A$
Background: Pathload
(Jain & Dovrolis, SIGCOMM 2002)

- Create short-lived congestion events
- send equi-spaced probes in a stream, detect one-way delay (OWD) trends using pair-wise comparison (PCT) and pair-wise difference (PDT) tests
- if probe stream causes congestion, then stream rate measured at receiver will be less than what was sent (OWD shows increasing trend)

\[
\frac{r_{in}}{r_{out}} = \begin{cases} 
\leq 1 & r_{in} \leq A \\
> 1 & r_{in} > A 
\end{cases}
\]

- Take \(N\) measurements with streams of length \(k\) to converge to an estimate of \(AB\)
Proposed calibration framework

• Experiments in a controlled laboratory environment — complete visibility
• Considerations for assessing measurement tools
  • Basic assumptions and models for tools and network
  • Limitations of host hardware and operating system
  • Network load (potential skewing) imposed by tools
  • Parameter selection and sensitivity
• Framework is not meant as a panacea
Calibration strategy

1. Design appropriate test environments where “ground truth” can be established
   - *In situ* environments generally not appropriate

2. Identify relevant test suites to test issues such as host/OS capabilities, loading effects, etc.
   - Real routing hardware and end systems are necessary

3. Apply flexible analysis and visualization techniques
   - We use scatter/phase plot representation for initial analysis (Tool: *SPLAT* — PAM 2006)
   - Plot spacing prior to cross traffic interaction versus spacing after cross traffic interaction

![Diagram showing egress and ingress spacings with interaction and non-interaction scenarios]
Testbed setup

- Testbed with commercial IP routers and switches, commodity workstations, and Endace DAG monitors
- Three topological variants
- Environment can be used by external researchers
Initial calibration experiments

- Target packet spacing is 80 microseconds for each tool
  - Spruce: 1200 packet pairs, each packet 1500 bytes
  - Pathload: 1200 probes of 1309 bytes
- Wide range of actual input spacings
- Bimodal output spacings caused by 50 Mb/s constant bit rate cross traffic
Initial results

• Absolute magnitude of error in spacings can be large, but

• Error has zero mean — about 20 packets is necessary to converge

• Probe spacings are both expanded and contracted

• Consistent with different hardware / OS setups

![Graphs showing relative frequency distributions of send and receive errors in microseconds.](image)
Calibrated Pathload (Yaz)

- Basic features of Yaz
  - Consider both expansion and contraction of spacings (Pathload only considers expansion)
  - Use mean input and output spacings, not any individual spacings

\[
|\bar{g}_{in} - \bar{g}_{out}| = \begin{cases} 
  \leq \zeta \times \bar{g}_{in} & r_{in} \leq A \\
  > \zeta \times \bar{g}_{in} & r_{in} > A.
\end{cases}
\]

- \(\zeta\) is a threshold design parameter
- Use exponentially-weighted moving average filter when reporting AB estimates
Comparison experiments

- Three traffic scenarios, three topological setups
- Use Endace DAG monitors to provide ground truth
  - Enables detailed analysis of probe streams to support fine-grained comparison aspect of calibration
- How often are Pathload, Spruce and Yaz within a 10% threshold of the true available bandwidth?
  - What causes inaccuracies?
Comparison results

comparison results

constant bit rate cross traffic

long-lived TCP (one way)

long-lived TCP (two way)

web-like traffic (topology 1)

web-like traffic (topology 2)

web-like traffic (topology 3)

dashed line represents 10% desired accuracy threshold
What causes inaccuracies?

• How much of an effect does stream contraction play?

• A significant fraction of streams are compressed — effect is ignored by both Pathload and Spruce

• How sensitive are PCT and PDT metrics of Pathload?

• Calculate PCT and PDT metrics prior to hitting congested queue

• For 50 packet streams, PCT value is close to 0.55 threshold (above which a OWD trend is inferred); PCT also high for 100 packet streams

<table>
<thead>
<tr>
<th>traffic scenario</th>
<th>fraction of compressed streams</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBR traffic</td>
<td>0.212</td>
</tr>
<tr>
<td>Long lived TCP (one-way)</td>
<td>0.077</td>
</tr>
<tr>
<td>Long-lived TCP (two-way)</td>
<td>0.260</td>
</tr>
<tr>
<td>Web traffic (topology 1)</td>
<td>0.233</td>
</tr>
<tr>
<td>Web traffic (topology 2)</td>
<td>0.220</td>
</tr>
<tr>
<td>Web traffic (topology 3)</td>
<td>0.219</td>
</tr>
</tbody>
</table>
Comparison results summary

- For a Spruce-like budget, Yaz provides better accuracy than Pathload
- An example of what might be expected as a result of calibration exercise

<table>
<thead>
<tr>
<th></th>
<th>estimates produced</th>
<th>latency (sec) μ (σ)</th>
<th>iterations per est. μ (σ)</th>
<th>mean pkts per estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>pathload (k=100)</td>
<td>96</td>
<td>17.7 (3.8)</td>
<td>8.4 (4.8)</td>
<td>10080</td>
</tr>
<tr>
<td>pathload (k=50)</td>
<td>97</td>
<td>17.6 (3.8)</td>
<td>8.8 (4.2)</td>
<td>5280</td>
</tr>
<tr>
<td>spruce</td>
<td>156</td>
<td>10.9 (0.9)</td>
<td>n/a</td>
<td>200</td>
</tr>
<tr>
<td>yaz</td>
<td>446</td>
<td>3.8 (1.5)</td>
<td>6.1 (8.8)</td>
<td>366</td>
</tr>
</tbody>
</table>
Summary and conclusions

• *In vitro*-like testing plays a pivotal role in tool calibration

• *Yaz* was created based on observations in such an environment

• Code will be made available: [http://wail.cs.wisc.edu](http://wail.cs.wisc.edu)

• *Splat* was also developed for scatter and phase plot visualization in this environment (PAM 2006)

• Future: generalize approach to more active measurement tools
the end

wail.cs.wisc.edu