Introduction to Networking and Systems Measurements

Measurement Pitfalls

Dr Noa Zilberman
noa.zilberman@cl.cam.ac.uk

Prof Andrew W. Moore
andrew.moore@cl.cam.ac.uk
Common Measurement Pitfalls

- What are the hidden assumptions?
- What did you not notice (in the system, setup, ….)?
- What can your tool do?
- Vantage points
- Repeatability pitfalls
- Performance pitfalls
- Reading the results
Hidden Assumptions - Examples

- The path from A to B is the same (reverse) as the path from B to A
- There is no packet reordering
- Device throughput is the same for all packet sizes
- Test packets will experience the same effects as application’s traffic
- The effect of DNS lookup is negligible
- The measurement tool has negligible overhead
- Previous work was correct
System and Setup
Did you notice that….

- There are other jobs running on the same core
- ICMP traffic is throttled by the OS
- CPU frequency scaling is enabled
- The CPU that you are using is not connected directly to the NIC
- Kernel version has been updated overnight
- The 2x40G NIC uses PCIe Gen 3 x8 (~60Gbps)
- There is a new Errata…
What can your tool do? - Examples

- SSD can write at 450MB/s
  - Don’t try to write data captured at 10Gbps
- The latency for reading CPU timestamp is ~tens of cycles
  - Don’t try to use it to measure cache access time
- DAG resolution is 4ns
  - Don’t try to measure the propagation delay through 1m fibre
- OSNT can only capture at low rate
  - Don’t try to measure latency of 10Gbps flows
Vantage Points: Example 2 (Lecture 5)

- Mellanox Spectrum vs Broadcom Tomahawk
  - Tolly report, 2016

- Bandwidth distribution, 3→1 scenario
  - Source ports 25,26,27, Destination port 31
    - 33% BW from each port, on both devices
  - Source ports 24,25,26, Destination port 31
    - 33% BW from each port, on Spectrum
    - 25% from ports 25,26, 50% from port 24 on Tomahawk

- What does it mean?
Repeatability Pitfalls - Examples

Running on bare metal (local)

Running on a VM (local)

Running in the cloud

Increased performance variance
Repeatability Pitfalls - Examples

Apache Webserver - Running in the cloud
38% difference in median performance
Latency Pitfalls - Examples

- What is the definition of “latency”?  
  - Propagation delay? Inter packet gap? Round trip time? Flow completion time?

- How was the latency measured?  
  - Start of packet to start of packet? Start of packet to end of packet?  

- Where was the timestamp taken?  
  - …and how did it affect the measurement?

- Resolution, precision and accuracy…
Bandwidth Pitfalls - Examples

- What is the definition of "bandwidth"?
  - Link capacity? Average throughput? Peak throughput?
- Controllability
  - Packet size? Protocol? QoS?
- What was the status of the network?
- Net neutrality?
- Did you pass through the bottlenecks?
- Resolution, precision and accuracy…
Example: Timestamp difference between ports

- Recall Lab 2, experiment 2.1 b
- Measuring the timestamp difference between 2 ports:
Example: Timestamp difference between ports

- 100,000 packets, \(1024\)B
- Different Inter Packet Gaps (IPG)
Example: Timestamp difference between ports

- 100,000 packets, 64B
- Different Inter Packet Gaps (IPG)
Example: Timestamp difference between ports

- 100,000 packets, **64B**, running 10 times
- Same Inter Packet Gap (IPG)
Example: Switch Throughput

- The reported iperf result for a NetFPGA reference switch is 9.4Gbps
- User complaint: I see only 8.9Gbps and packet drop in the switch
Example: Switch Throughput

- Debug: Have you tried changing rx-usec?
- User: no more packet drop in the switch!
- …but bandwidth is down to 7.5Gbps…

- New insight: NIC used on reference setup (Solarflare) is different than the NIC used by user (Intel)
- (skipping a few steps forward)
Example: Switch Throughput

- Switch throughput over time (10ms sampling resolution)
Example: Switch Throughput

- Switch throughput over time (100µs sampling resolution)
Example: Switch Throughput

- What else is different?
Example: TSC Access

- **Goals:**
  - Evaluate the accuracy & precision of time-taking using CPU time stamp counter (TSC)

- **Methodology:**
  - Read TSC twice
  - Measure the time-gap between the two consecutive reads

- **Results:**
  - Min/Median/99.9%: 9ns/10ns/11ns
Example: TSC Access

```c
while (!done)
{
    // Read TSC twice, one immediately after the other
    do_rdtscp(tsc, cpu);
    do_rdtscp(tsc2, cpu2);
    // If the gap between the two reads is above a certain threshold, save it
    if (((tsc2 - tsc > threshold) && (cpu == cpu2))
        buffer[samples++] = tsc2 - tsc;
}
```
Example: TSC Access

What happens over time?

![Graph showing gap over time for Console and SSH access.](image-url)
Example: TSC Access

- Source data:

<table>
<thead>
<tr>
<th>( X \leq )</th>
<th>User space Events</th>
<th>( X \leq )</th>
<th>Kernel Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>91428291492</td>
<td>9</td>
<td>11117819727</td>
</tr>
<tr>
<td>11</td>
<td>404700</td>
<td>10</td>
<td>3973891503</td>
</tr>
<tr>
<td>12</td>
<td>268521</td>
<td>49</td>
<td>287</td>
</tr>
<tr>
<td>22</td>
<td>268291</td>
<td>53</td>
<td>201</td>
</tr>
<tr>
<td>120</td>
<td>267465</td>
<td>98</td>
<td>90</td>
</tr>
<tr>
<td>1097</td>
<td>10768</td>
<td>1155</td>
<td>86</td>
</tr>
<tr>
<td>10869</td>
<td>1</td>
<td>1184</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1241</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1982</td>
<td>1</td>
</tr>
</tbody>
</table>
Example: TSC Access
Validation

- Measurements need to be validated
- Don’t make assertions!
- Use ground truth (where available)
- Compare different tools and methodologies
- Do the results make sense?
  - RTT can’t be faster than traveling at the speed of light…
- Have I mentioned validation?
Labs 4-5 and Final Report

- Each student assigned an artifact
- Black-box evaluation
- Running in Azure – Accept invite!
- Read the handout before Lab 4
  - Lab will provide a walk through and Q&A time
- Prepare a reproducibility review of a paper
- Extra mark for reproducing an experiment from the paper
- Lab 5 – discussion of the test plan, Q&A
Final Report - Recommendations

- Include all figures within the report
  - Use proper scale, adapt the template if need be
- Make sure that your environment does not affect the results
- Do not make assertions
  - Support your claims through experimentations
- Discuss your results in depth:
  - Compare and contrast results gained through different vantage points, using different tools, on different platforms etc
  - Provide side-by-side comparisons
  - Use the questions in the handouts as guiding examples
- Use the right terminology (accuracy, precision, resolution)
- Correct typos and grammar mistakes
- Make sure not to run out of budget
- Follow the instructions in the handout
Course Summary

- This course has covered measurements tools and measurement techniques
- But also “why out most basic assumptions are wrong”, “graphs lie”, “what you don’t know about your system”, …
- Remember:
  - Constant vigilance
  - Look at the data, best-practice, think.
- Applies to all types of measurements