A Low-Latency Network Monitoring Platform

Jeff Rasley (jeffra@cs.brown.edu), Brent Stephens (brents@rice.edu),
Colin Dixon, Eric Rozner, Wes Felter, Kanak Agarwal, John Carter, and Rodrigo Fonseca

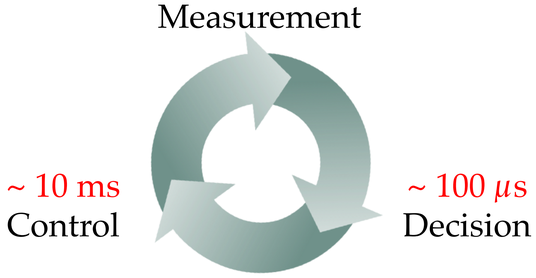
BROWN

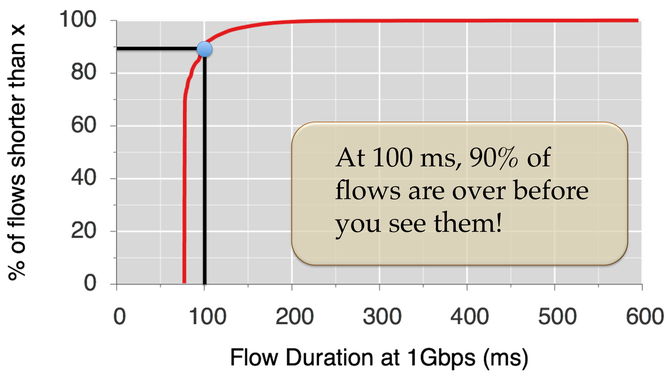
IBM Research
Austin Research Lab

RICE

Measurement is Slow

- SDNs introduce the possibility of self-tuning networks that can quickly detect and react to congestion or failures.
- As link speeds increase to 10 Gbps+ existing control loops must decrease.





Background TCP flows, Microsoft data center from DCTCP.

Alizadeh et al. Sigcomm '10

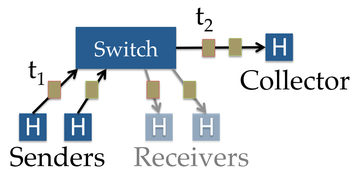
Summary

- Problem: Network monitoring via port counters or sFlow takes between 100 ms–1 sec+ to obtain flow information.
- Solution: Mirroring all traffic through a single mirror port on each switch lowers measurement latency by **3–4 orders of magnitude!** This give us latencies of 250μs–6.5ms vs. 100ms–5s for the state-of-the-art.
- How: Does this actually work? Ask for more details!

What Can Go Wrong?

- Packets don't include input/output port information
  - ✓ Share switch state with collectors
- Unknown drop policy when mirror port fills
  - ✓ Compute rate estimates via TCP sequence numbers
- Oversubscribed port may occupy switch buffer space
  - ✓ True, but negligible effect

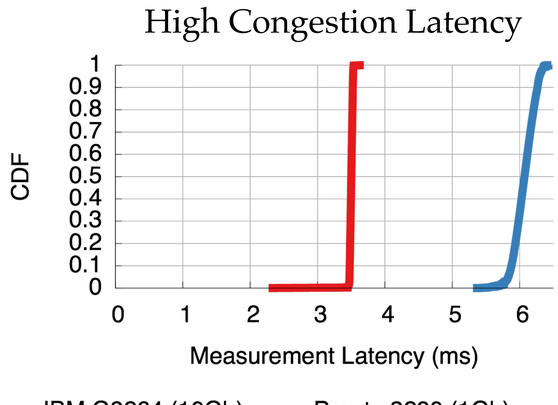
Sample Latency



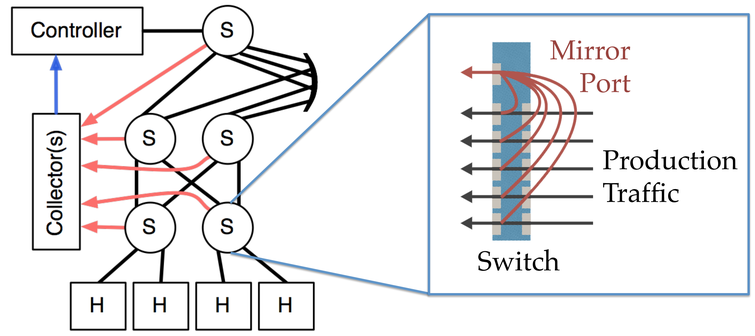
Latency = $t_2 - t_1$

Low Congestion Latency:

75–150 μs

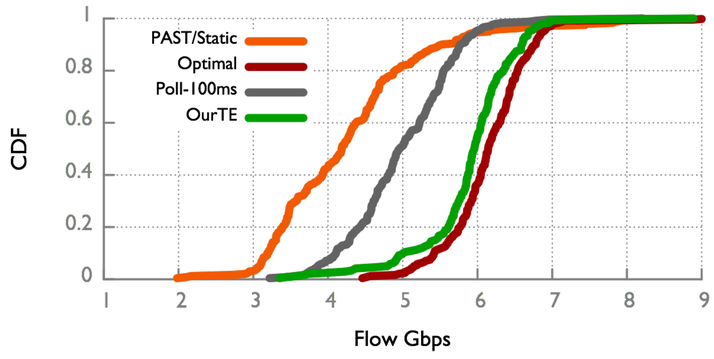


Architecture



- Mirror all production traffic to a single oversubscribed mirror port across all switches in the network
- Process packets with collector machines using netmap/DPDK

TE within 4% of Optimal



Stride(8) 100 MiB Workload CDF of Flow Throughput