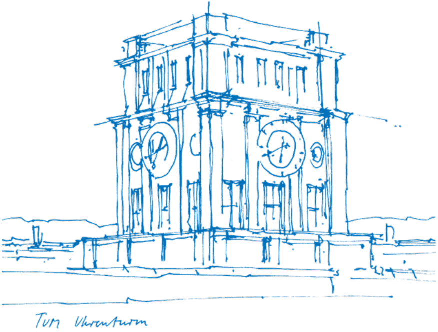
Chair of Network Architectures and Services
Department of Informatics
Technical University of Munich

Revisiting Benchmarking Methodology for Interconnect Devices

Daniel Raumer, Sebastian Gallemüller, Florian Wohlfart,
Paul Emmerich, Patrick Werneck, and Georg Carle

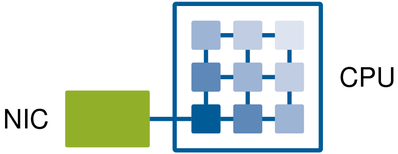
July 16, 2016



Contents

Why to revisit benchmarking state of the art?

- Numerous standards, recommendations, best practices
  - Well-known benchmarking definition RFC 2544
  - Various extensions
  - Divergence of benchmarks
- New class of devices
  - High speed network IO frameworks
  - Virtual switching
  - Many core CPU architectures:

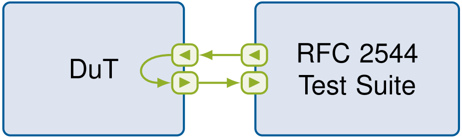


Common metrics

- Throughput: highest rate that the devices under test (DuT) can serve without loss.
- Back-to-Back frame burst size: longest duration (in frames) without loss.
- Frame loss rate: percentage of dropped frames under a given load.
- Latency: average duration a packet stays within the DuT.

- … extended metrics, e.g., FIB-dependent performance
- … additional SHOULDs, rarely measured

Case study: RFC 2544 benchmarks

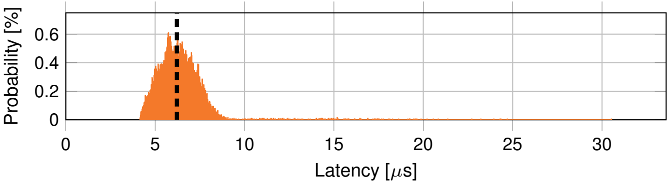


Three different DuTs

- Linux router
- FreeBSD router
- MikroTik router

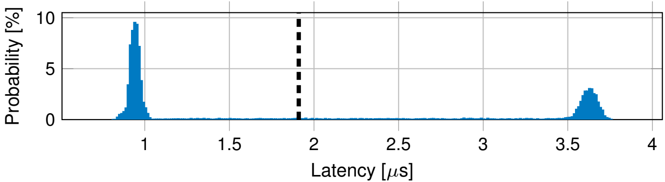
Flaws of benchmarks: selected examples

Meaningful latency measurements: case study



- FreeBSD, 64-byte packets
- Average does not reflect long tail distribution

Meaningful latency measurements: 2nd example

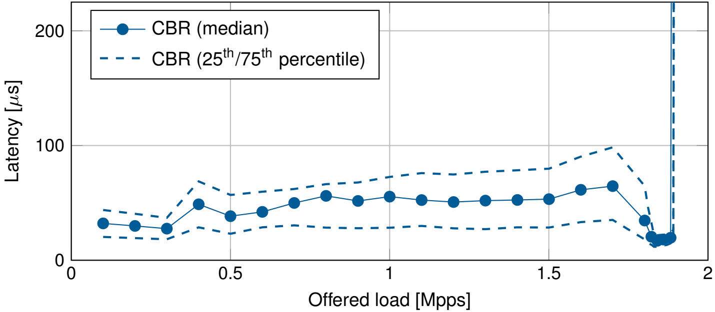


- Pica8 switch tested in [IFIP NETWORKING 16]
- Different processing paths through a device
- Bimodal distribution
- Average latency is misleading

→ Extensive reports: histograms for visualization

→ Short reports: percentiles (25th, 50th, 75th, 95th, 99th, and 99.9th)

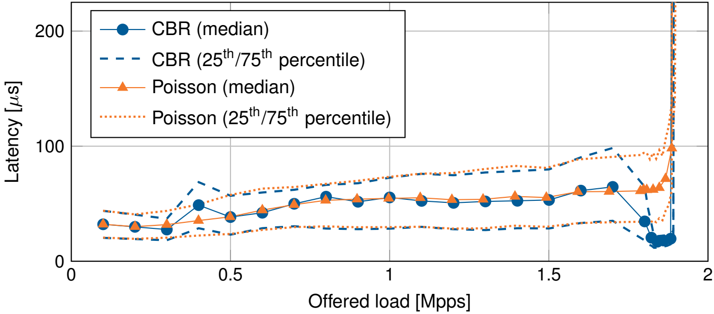
Latency under load



- Open vSwitch (Linux NAPI & ixgbe) [IMC15]
- Latency at maximum throughput is not worst case

→ Measurements at different loads (10, 20, ..., 100% max. throughput)

Traffic pattern & latency

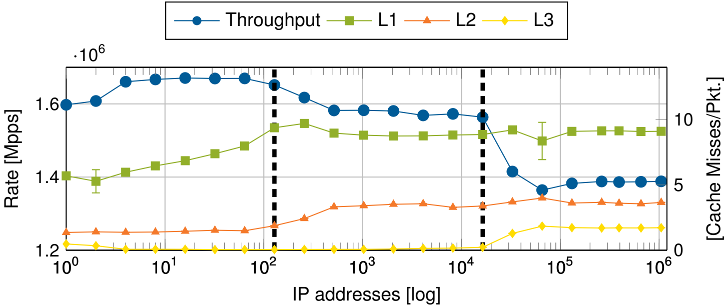


- Open vSwitch (NAPI + ixgbe) [IMC15]
- Different behavior for different traffic patterns

→ Tests with different traffic patterns

→ Poisson process to approximate real world traffic

Omitted tests



- CPU caches affect the performance

→ Additional tests for certain device classes

→ Functionality dependent tests

Reproducibility of configurations

- Manual device configuration is error prone
- Device configuration is hard to reproduce

→ Reproducible configuration of DuT via scripts

→ Configuration scripts executed by benchmarking tool

Conclusion

- Novel class of devices requires additional tests
- There are arguments for reconsidering best practice:
  - Average latency may be misleading
    $\rightarrow$ Histograms / percentiles
  - Latency is load dependent
    $\rightarrow$ Measure 10, 20, ..., 100% of max. throughput
  - CBR traffic is a unrealistic test pattern
    $\rightarrow$ Poisson process
  - Device specific functionality
    $\rightarrow$ Perform device specific benchmarks
  - Manual configuration is error prone
    $\rightarrow$ Automatic configuration by benchmark tool

Novelty: RFC 2544 test suite on commodity hardware

- MoonGen [IMC15] is a fast software packet generator
- Hardware-assisted latency measurements (misusing PTP support)
- Precise software rate control and traffic patterns

- http://net.in.tum.de/pub/router-benchmarking/
- RFC 2544 benchmark reports for Linux, FreeBSD, and MikroTik
- Early version of the MoonGen RFC 2544 module