OCP SUMMIT

March 20-21
2018
San Jose, CA

OPEN. FOR BUSINESS.

Using SmartNICs to Reduce Server Latency

Nic Viljoen, Director-Software Engineering, Netronome
nick.viljoen@netronome.com

Sections

NETRONOME

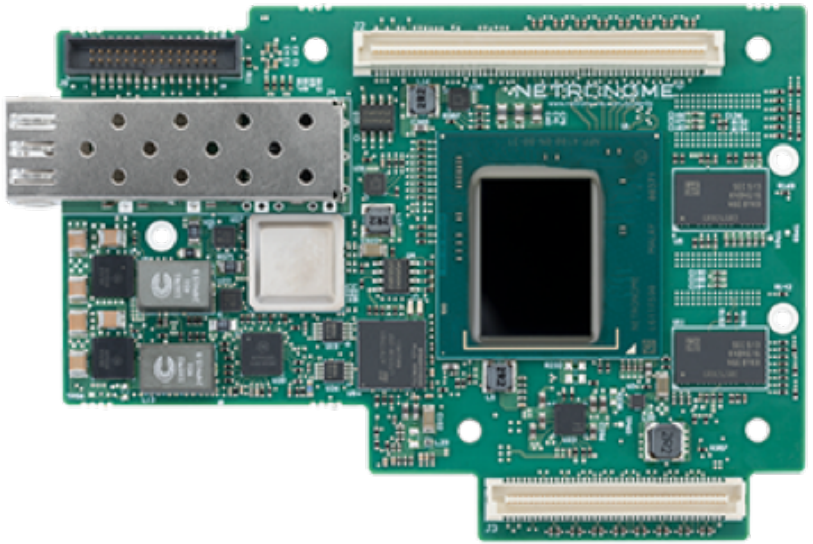
- Intro
  - Netronome and OCP
  - What's a SmartNIC?
- Why is Latency a problem?
- How can SmartNICs Help?
  - Single-Host
  - Multi-Host
- Summary

© 2018 NETRONOME SYSTEMS, INC.

Netronome and OCP

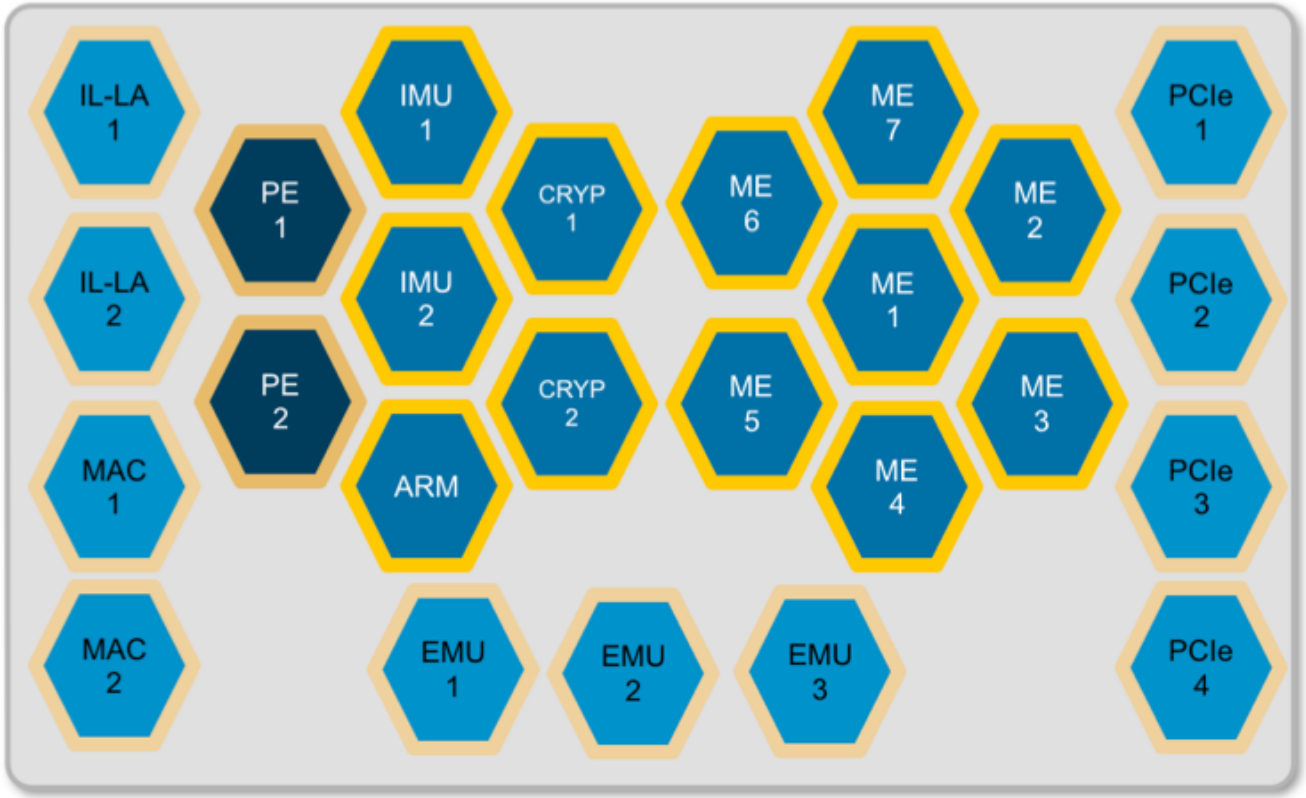
NETRONOME

- First 25/50G SmartNIC on OCP Mezz v2
  - Fully programmable-72 cores, 8 threads per core
- 15-25W (Depending on use case)
- Contributing to OCP Mezz v3
  - Jack Dawson
  - john.dawson@netronome.com



© 2018 NETRONOME SYSTEMS, INC.

SoC Architecture



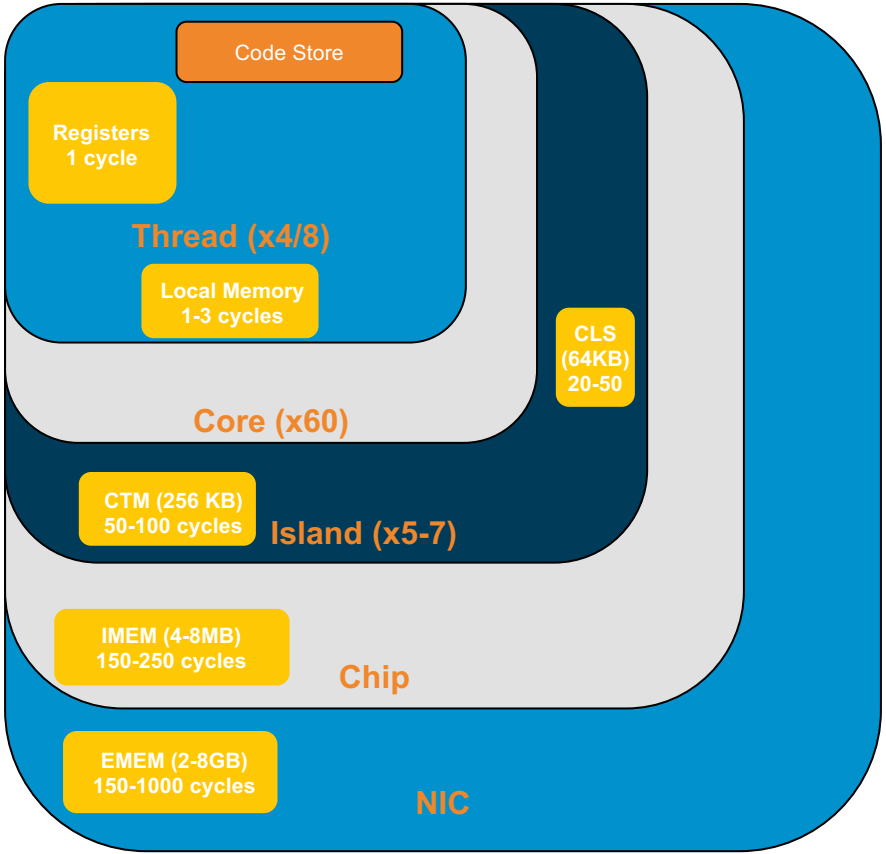
- High Performance Interconnect – DSF
- Modular Island based architecture
- Scalable with process node
- Capability to add/ remove islands based on customer requirements
- Low latency deterministic paths between islands

© 2018 NETRONOME SYSTEMS, INC.

Transactional Memory Architecture

NETRONOME

Latency hidden by co-operative multi-threading



© 2018 NETRONOME SYSTEMS, INC.

1. Why is Latency a Problem?

© 2018 NETRONOME SYSTEMS, INC.

Lossy networks need all the help they can get

NETRONOME

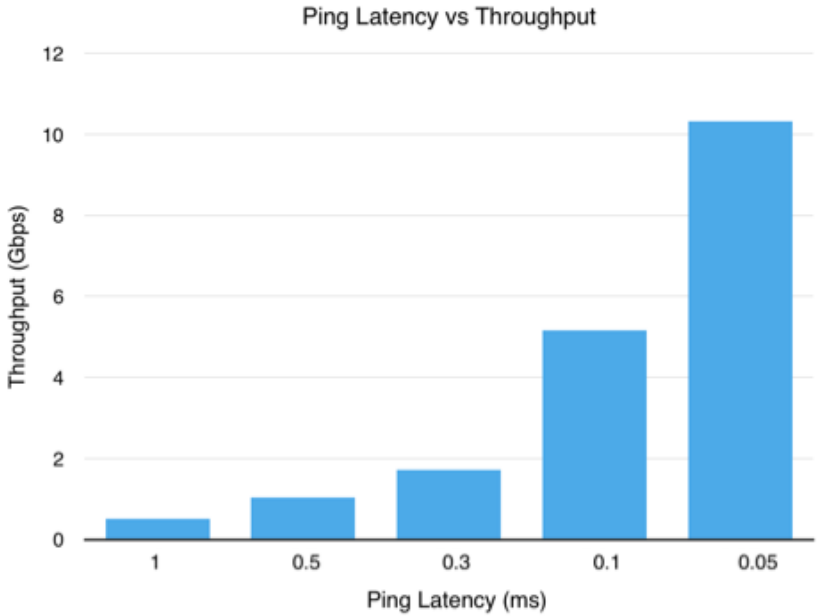
- Focus of this talk is TCP (and a small bit of UDP)-nothing fancy
- Latency affects
  - Throughput
  - Remote procedure call reliability
  - Network hygiene
- Simple but useful tool
  - https://wand.net.nz/~perry/max_download.php

© 2018 NETRONOME SYSTEMS, INC.

Window size, losses and latency (TCP 101)

NETRONOME

- Going from 1ms to 0.05ms increases throughput by about **~20x**
  - **Window size-why does it really matter?**
- Any losses will accentuate this
- **Reducing latency increases robustness**



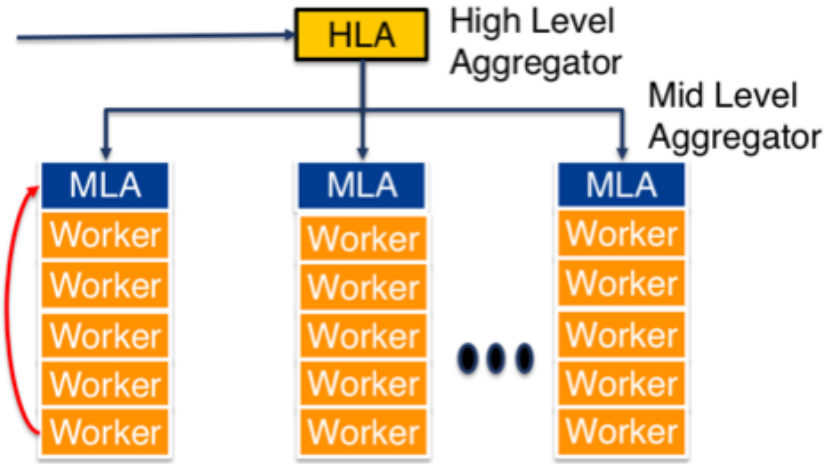
© 2018 NETRONOME SYSTEMS, INC.

9

Remote Procedure calls

NETRONOME

- Web users leave websites if interactions have too much latency
- **This is decreasing with the advent of VR/AR**
- **Tail latency is key**
  - **1/100 workers exceeds P99 latency budget for process**



© 2018 NETRONOME SYSTEMS, INC.

2. How Can SmartNICs Help?

© 2018 NETRONOME SYSTEMS, INC.

Processing on the NIC

NETRONOME

- Processing certain types of packets on the NIC significantly reduces latency
- **Want to be able to run own applications**
- Don’t want to leave upstream (Linux)
- How can custom datapath offload be achieved while staying within upstream?
  - eBPF

© 2018 NETRONOME SYSTEMS, INC.

eBPF

NETRONOME

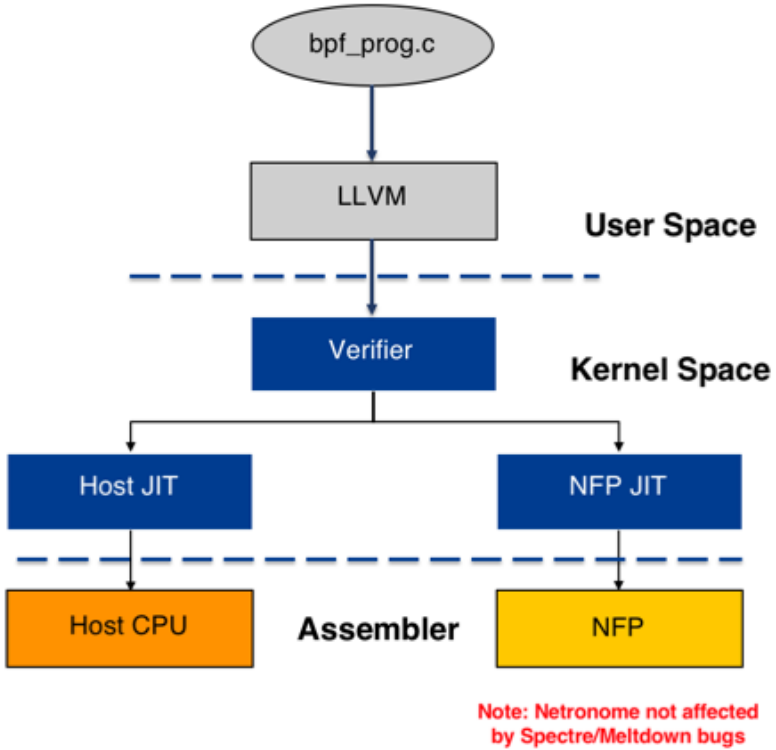
- Small kernel-based virtual machine
  - 10 64-bit registers
  - 512 byte stack
  - Max 4k RISC bytecode instructions
  - Infinite size key-value stores (maps)
- BPF has a verifier to ensure programs do not contain non-permitted state
- Helpers do essential work outside of BPF (e.g map lookups, header extend)

© 2018 NETRONOME SYSTEMS, INC.

eBPF Offload

NETRONOME

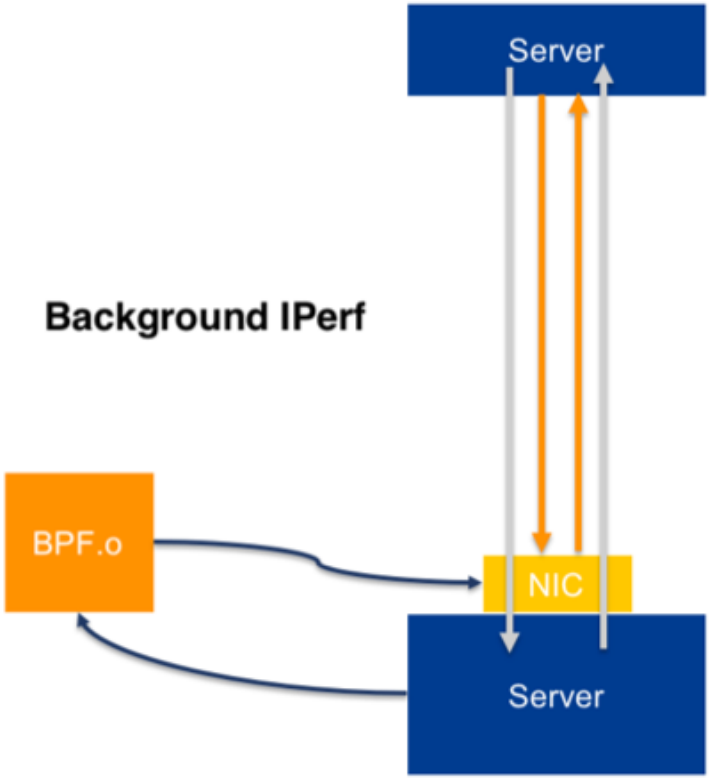
- What is eBPF?
  - Small kernel-based virtual machine
  - Compiled from C/Go/Rust/P4 by LLVM
  - Verified and JITed by kernel
- Why eBPF?
  - Emerging technology in kernel
  - Used by Facebook, Cloudflare, many others
  - BPFilter is key new firewall method
- eBPF Offload
  - Transparently offload XDP and cls_bpf (TC)
  - Means NFP can immediately offload new kernel innovations



© 2018 NETRONOME SYSTEMS, INC.

Ping Latency Example

NETRONOME



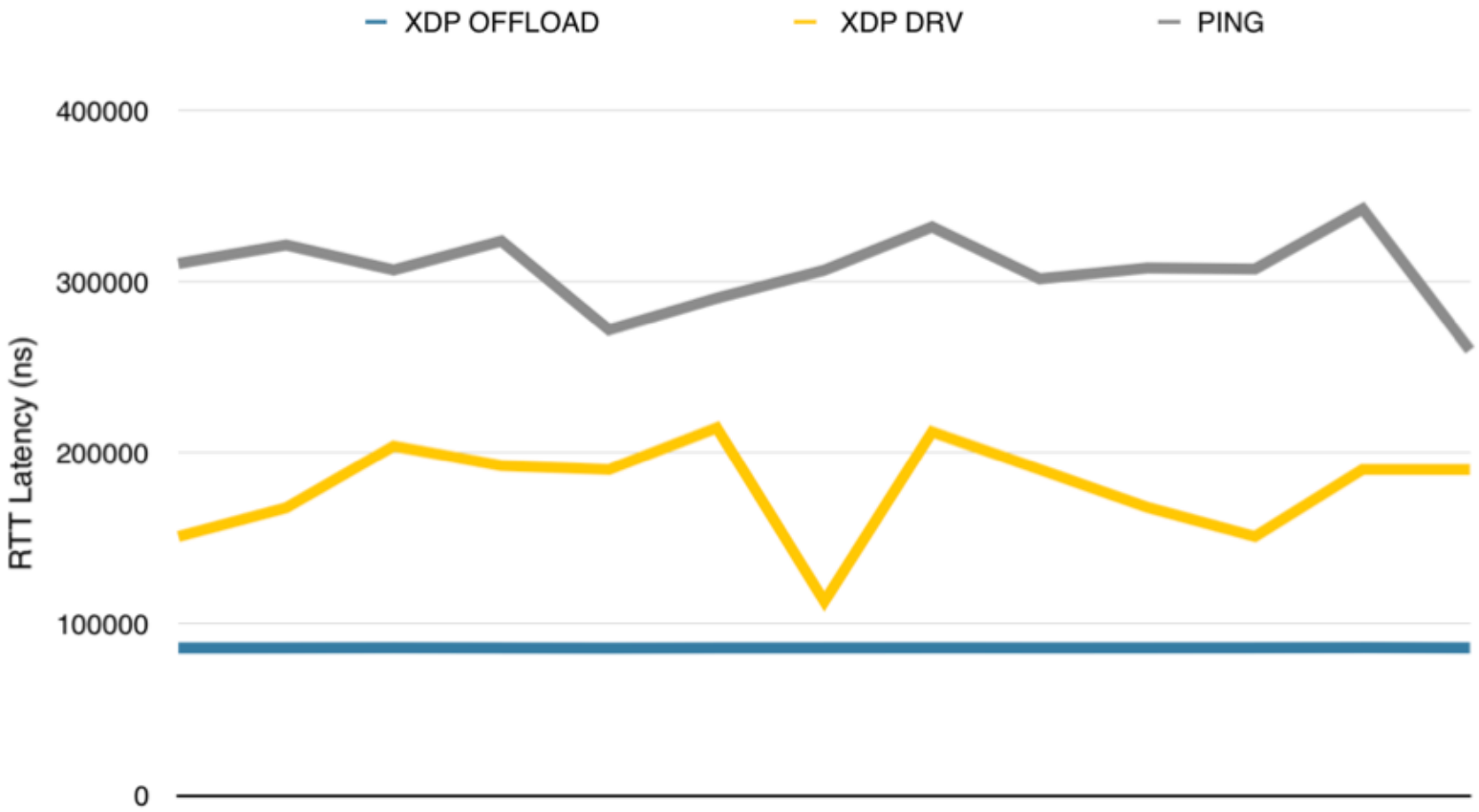
Compare RTT-**Remember the graph on slide 10**

DRV and PING done with multiple NICs

© 2018 NETRONOME SYSTEMS, INC.

Ping Latency Example

NETRONOME

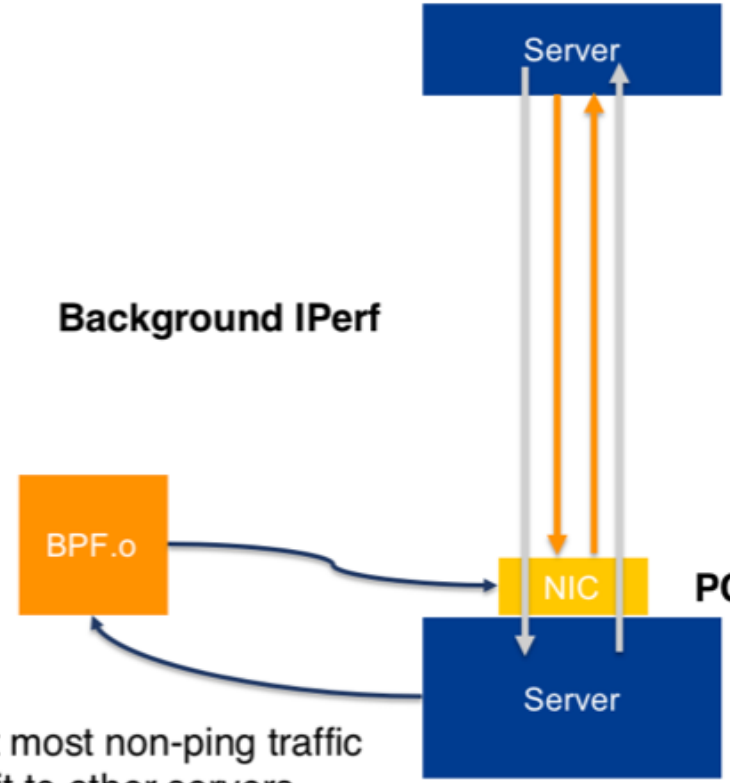


© 2018 NETRONOME SYSTEMS, INC.

Ping Latency + Filter (Multi-host)

NETRONOME

There are other methods to reduce MH latency with a SmartNIC.
Some are much more general



Compare RTT-**Remember the graph on slide 10**

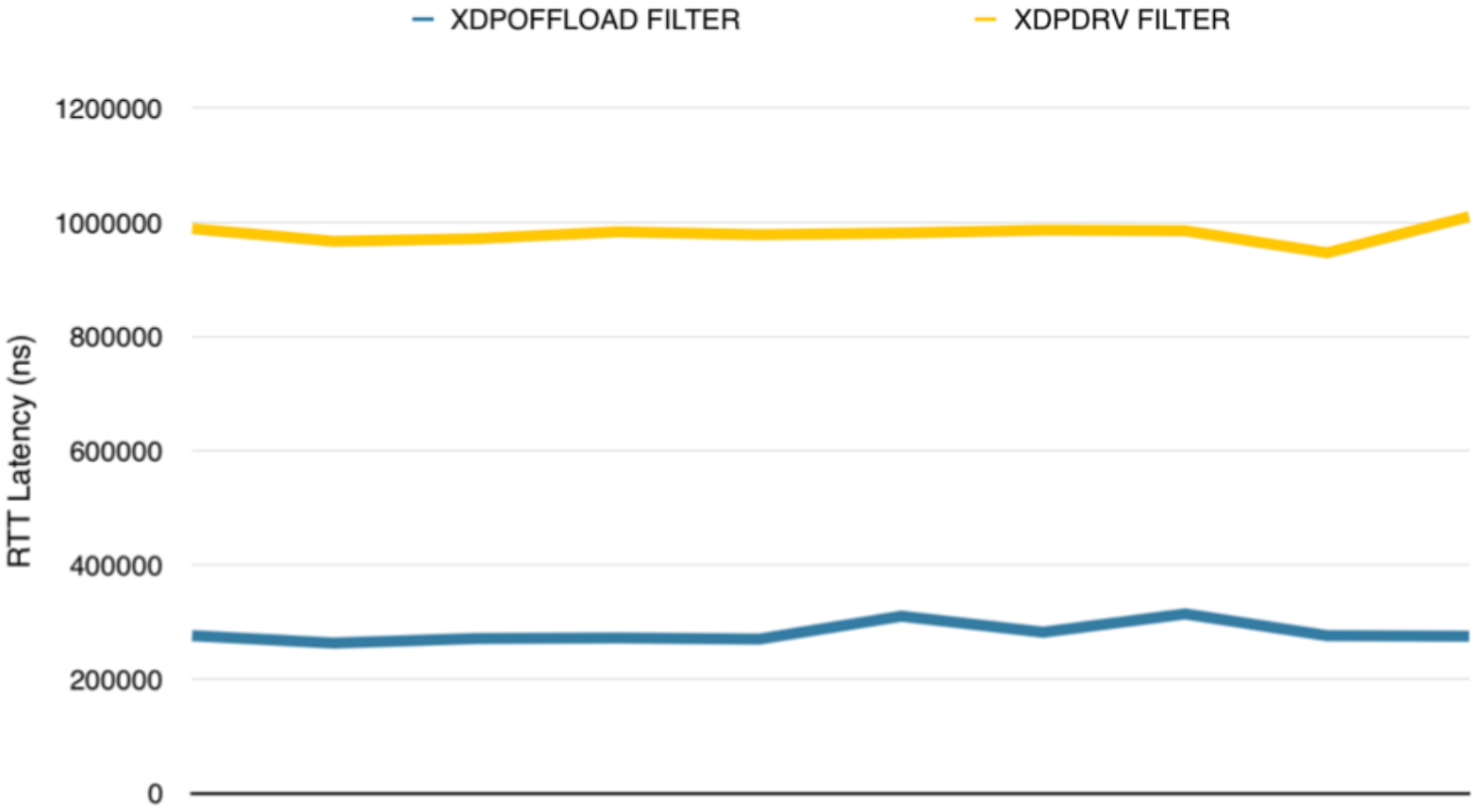
PCIe Limited to x4/x2-can cause drops

BPF program filters out most non-ping traffic and load balances it to other servers

© 2018 NETRONOME SYSTEMS, INC.

Latency (Multi-host)

NETRONOME



© 2018 NETRONOME SYSTEMS, INC.

Summary-How does this help ocp?

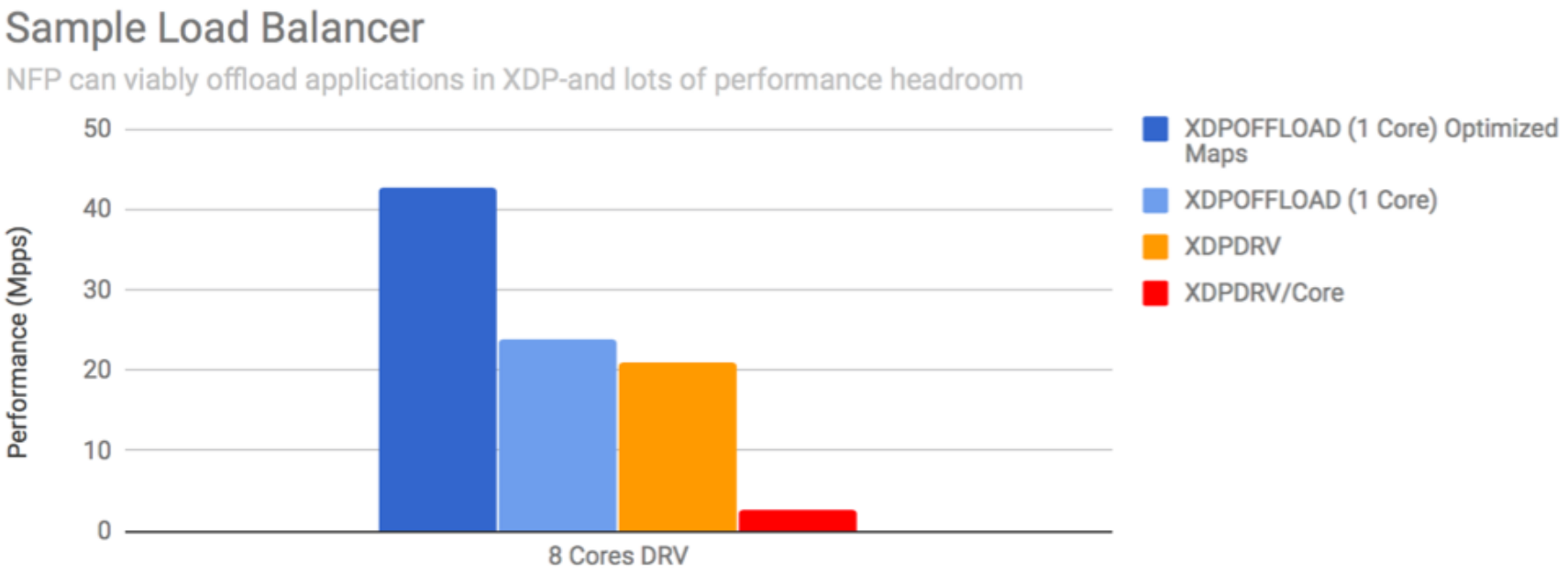
NETRONOME

- The crossing point between HW and software is hard
- But the rewards can be very interesting
- How can NIC level programmability become more tightly entwined with OCP?

© 2018 NETRONOME SYSTEMS, INC.

Load Balancer Throughput

NETRONOME



© 2018 NETRONOME SYSTEMS, INC.

OCP SUMMIT