SNIA DSI '16 DATA STORAGE INNOVATION CONFERENCE
Innovation in Storage Products, Services, and Solutions
June 13-15, 2016 | Marriott San Mateo | San Mateo, CA

Application Advantages of NVMe over Fabrics

RDMA and Fibre Channel

Brandon Hoff

Broadcom Limited

Tuesday, June 14 2016

10:55 – 11:35 a.m.

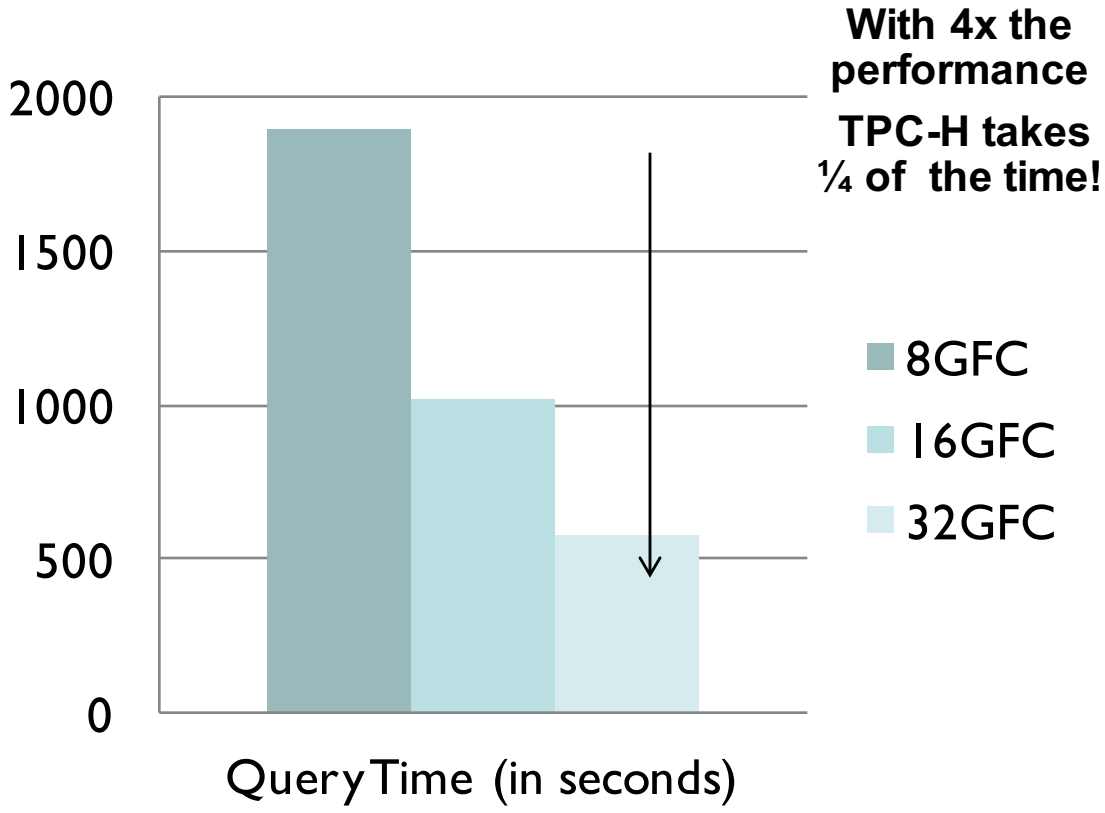
Agenda

- Applications that have a ‘need for speed’
- The Benefits of NVMe
- The need for NVMe Fabrics
- NVMf
  - The Standard
  - Code Submission to the community
- Broadcom Reference Hardware Architecture for SDS
- Storage Array Requirements
- NVMe over RDMA and NVMe over Fibre Channel
- Closing

Testing the New Paradigm

SQL TPC-H Test Results by Demartek

- Performance hungry applications can consume all the performance that we can give them
- The race is on to improve performance and reduce latency



Source: Demartek, February 2016

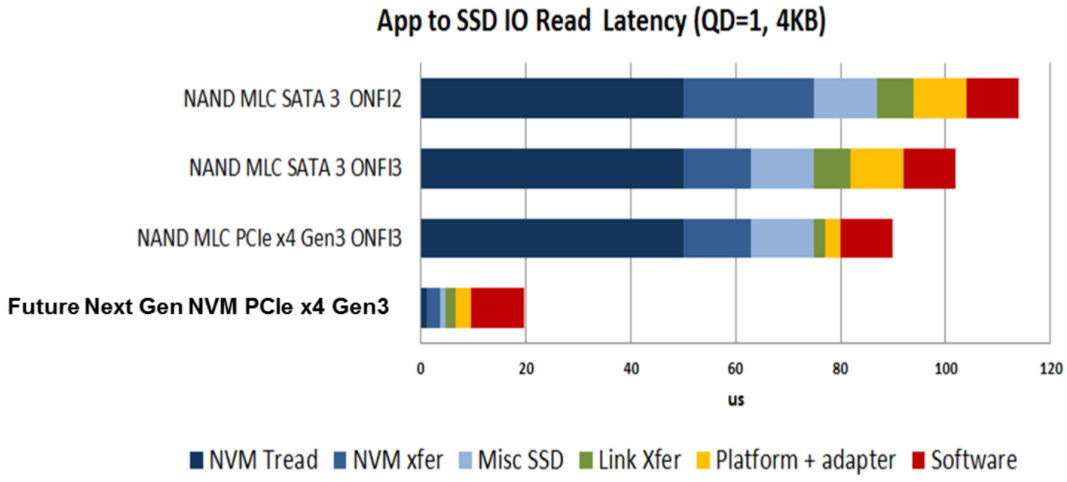
http://www.demartek.com/Demartek_Emulex_LPe32000_Gen6_FC_Evaluation_2016-03.html

In an All-Flash World – Performance Matters!

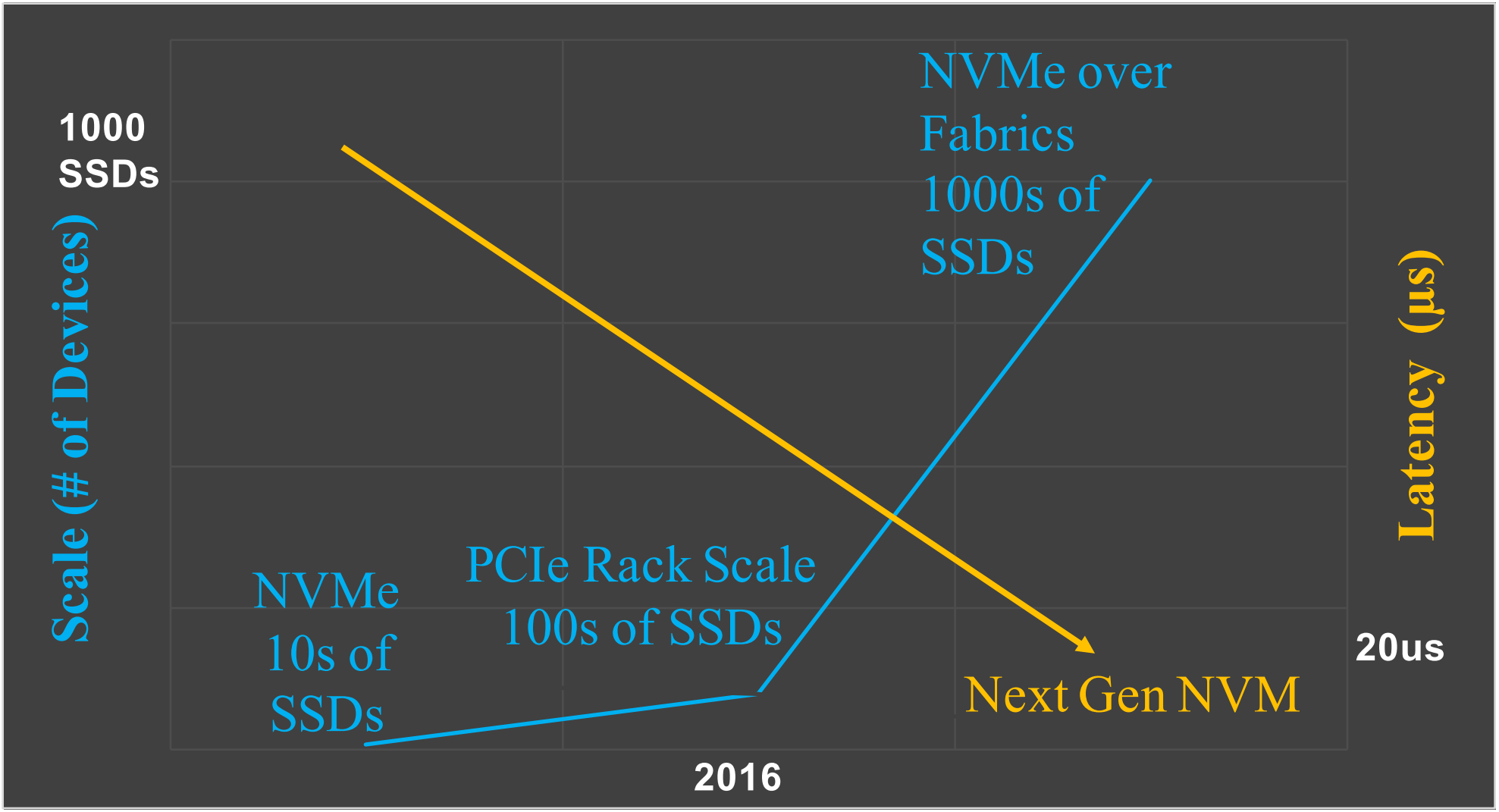
The NVMe Opportunity

- The industry is very focused on improving storage performance with NVMe
- The bottleneck is moving from the storage device to the software & network stack
  - Low NVMe latencies expose new bottlenecks
- Next step is to scale the number of devices beyond a rack

Storage latencies are dropping by orders of magnitude

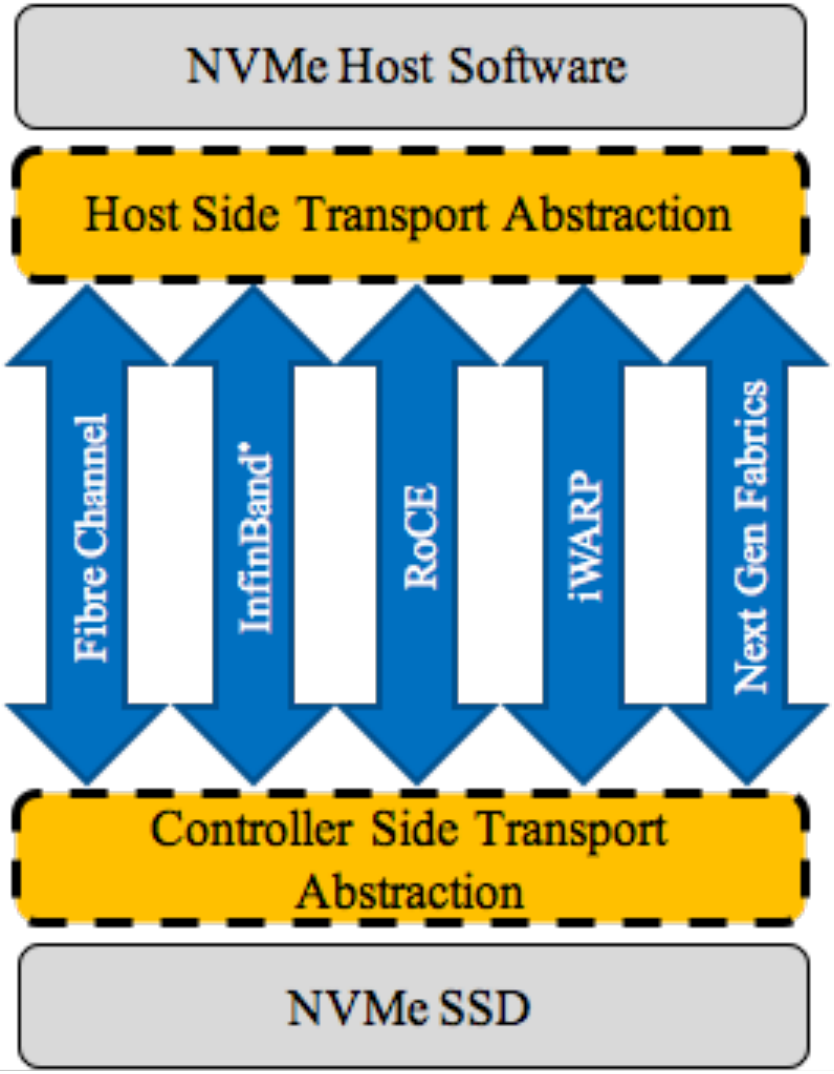


NVMf Evolution



NVMe for Fabrics
NVM Express, Inc.

- Industry-wide NVMe community defining NVMe over Fabric
- Dozens of participating companies
- Working group actively reviewing final items in draft spec now
- NVMe Fabrics Linux Driver working group active



NVMe over Fabrics

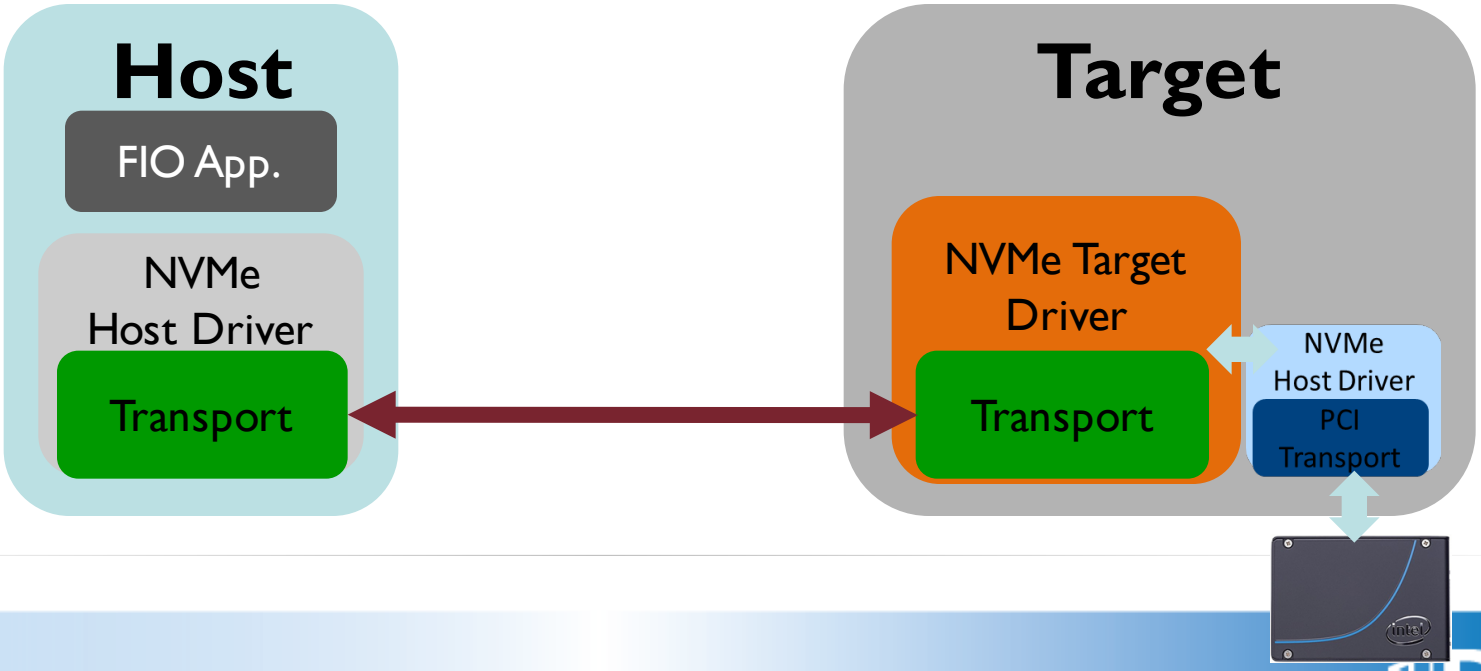
- Why is NVMe Fabrics important?
  - Simple, low-latency transport leveraging NVMe
  - Layered architecture for scale-out over any fabric
- When is it happening?
  - NVMe over Fabrics Specification Released June 6th 2016
  - NVMe fabrics software stack contributed to the community
  - Multiple prototypes now, early releases in 2016
- What can we build with it?
  - Fabric-connected “JBOD” of NVMe devices
  - Low-latency NVMe end-to-end storage servers
  - Fastest all-Flash Arrays and NG-NVM storage arrays

NVMe over Fabrics Stack Development

- Host and Target source code in a public repository on gitlab.com
  - Developers create git patches and email them to the WG reflector
  - WG members approve and/or comment on the new patches
  - Maintainer integrates approved patches into the for-next branch
- Multiple members actively submitting patches
  - Adding new functionality based on latest fabrics TP specification
  - Fixing bugs that have been identified during testing
  - Several rebases to latest upstream Linux kernel functionality

Linux Fabric Driver Architecture

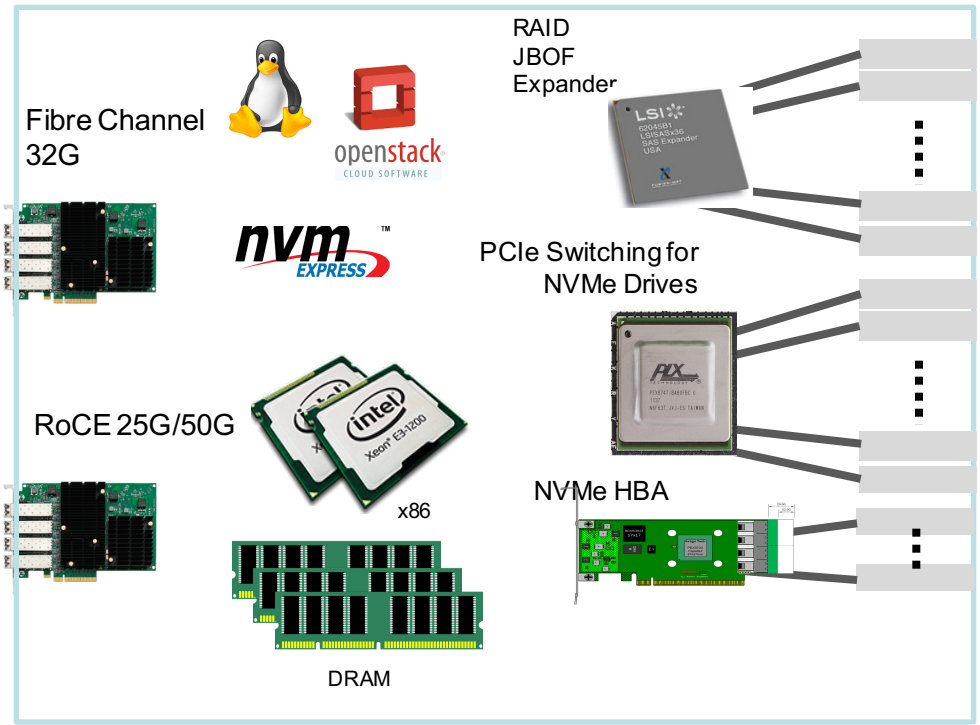
- Initiator and target software stacks for NVMe over Fabrics
  - RDMA and Fibre Channel supported, initially
- Standardized Latency Test
  - Determine latency to access an NVMe PCIe SSD remotely vs locally



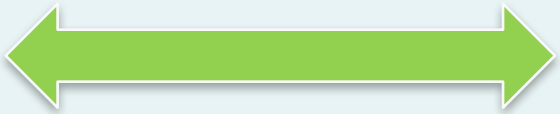
Broadcom SDS Reference Platform

Enterprise Proven: Performance and 74% lower TCO for NVMe Drives

- Broadcom reference design
  - Broadcom Dual-Mode Fibre Channel (NVMe and SCSI)
  - Broadcom 25G RoCE NIC
  - RAID/JBOD – Broadcom SAS/SATA/NVMe
  - Broadcom DCSG PCIe switching
  - Broadcom DSD NVMe ASICs (on the NVMe SSD)
- Hardware Platform for
  - Traditional customers – Build out and test performance benefits
  - Ready to go Hardware Platform for Startups – Predefined and configured hardware system
- High Performance Architecture
  - High performance data protection
  - Separate control and data paths
  - NVMe technology for low latency



Storage Array Performance Considerations

Standard Storage Arrays	NVMe Storage Arrays	NVMe JBOF
~200us Latency	~100us Latency	~40us Latency
Full Featured	Tradeoff enterprise features for performance	No Enterprise Features
Applications Trading, Financial, etc.		Applications Ceph, Hadoop, etc.

Different applications will require different storage architectures.

NVMe over RDMA (RoCE) Vs. NVMe over Fibre Chanel

	NVMe over RDMA (RoCE)	NVMe over FC
Throughput	1x at 25G, 4x at 100G	1x at 32G, 4x at 128G
Switch QoS	PFC	Credit based flow control
Kernel Bypass	Yes	Yes
Low Latency	Yes	Yes
CPU Overhead	Low	Low
Converged Ethernet	Yes	FCoE
Security	New	Proven

Different applications will require different network architectures.

Next Steps

- Understand the importance of latency in your applications
- Choose the network that fits your business needs
  - The Performance is identical between them
- Leverage the reference platforms for storage software development
- Applications will consume all of the performance we can give them.