



NVM Express Introduction & Tutorial

David Akerson, Intel Corporation

#OFADevWorkshop



“If I had asked people what they wanted, they would have said faster horses.”
- Henry Ford

What is *nvm*[™] EXPRESS ?



The industry standard for Enterprise and Client Non-volatile Memory based storage solutions

Standardizes register set, feature set, and command set to deliver performance

Architected from the ground up for Non-volatile memory to be more efficient, scalable, and manageable

Developed by an open industry consortium for Client, Enterprise and Data Center

NVM Express Community



NVM Express, Inc.

Consists of more than 80 companies from across the industry



Promoter Group

Led by a 13 company Board of Directors

Technical Workgroup

NVM Express Specifications - Queuing interface, NVMe I/O and Admin command set

Marketing Workgroup

NVM Express Awareness

Management Interface Workgroup

Out-of-band management over PCIe VDM and SMBus

NVM Express over Fabrics Workgroup

A flexible transport abstraction layer useful for many different fabrics



2014: An Amazing Year

JULY 18TH, 2014 by Josh Linden

Samsung Announces Industry's First 2.5-Inch NVMe SSD

STORAGE REVIEW.COM

Tweet 4
Share

Samsung has announced the XS1715, a 2.5-inch Non-Volatile Memory Express (NVMe Express) PCIe SSD. According to Samsung, the 1.6TB SFF-8639 NVMe SSD provides a sequential read speed at 3,000MB/s, six times faster than the company's current high-end enterprise SSD. The XS1715's random read performance is specified at up to 740,000 IOPS, more than 10 times as fast as existing SSD options.



REUTERS EDITION: U.S. SIGN IN

HOME BUSINESS MARKETS WORLD POLITICS TECH OPINION BREAKINGVIEWS

“...our first look at an NVMe drive, and the results are impressive.”

AnandTech

Specification
Client
Solid-State

JUNE 3RD, 2014 by Lyle Smith

Intel PCIe with NVMe

Intel has announced a new line of performance gains over SATA SSDs to meet the increasing needs for high cost of ownership. The new line includes workloads, and the Intel SSD DC P3700.



SEPTEMBER 3RD, 2014 by Adam Armstrong

Intel Announces New Series of NVMe-compliant Ultrastar SN100 PCIe SSDs

Intel has announced its new series of NVMe-compliant Ultrastar SN100 PCIe SSDs, the SN150 (3200GB and 1600GB) and the SN100 2.5-inch form factor (3200GB, 1600GB and 800GB) which integrates Toshiba's current MLC NAND flash. HGST indicates that it will drive the next wave of application acceleration with simplified PCIe SSD system integration coupled with HGST's experience in industry. The standardized NVMe protocol for PCIe SSDs enables broad system interoperability as well as ease of deployment, which results in lower cost of ownership.

Tweet 6
Share 5

ANNOUNCING



ULTRA SN100 NVMe
Accelerating performance

1.6 & 3.2TB capacities
HHHL & 2.5in form factors

SEPTEMBER 3RD, 2014 by Adam Armstrong

NVM Express Initiates NVM Express Over Fabrics

NVM Express Inc. announced today that it is initiating an effort to specify a standard for NVM Express (NVMe) over Fabrics. NVM Express will extend the benefits of NVMe to where usage over fabric is more appropriate than using PCIe. NVM Express also announced a revision of version 1.2 of its NVM Express Specification.

NVM Express is a scalable host controller interface designed for enterprise, data center, and client systems. NVMe has been utilizing PCIe interface based on SSDs. With NVMe over Fabrics will enable end-users to achieve faster application response times and better scalability across virtual data centers by connecting remote subsystems with flash appliances. Though this ability already exists, using NVMe end-to-end eliminates the increased latency through SCSI translation. End-users will see similar latency regardless if the SSDs are remote or local. NVMe over Fabric is aimed at flash appliances that use fabrics such as Ethernet with RDMA, InfiniBand, and Intel Omni Scale Fabric.

STORAGE REVIEW.COM

Tweet 7
Share 1

NVMe™ Driver Ecosystem



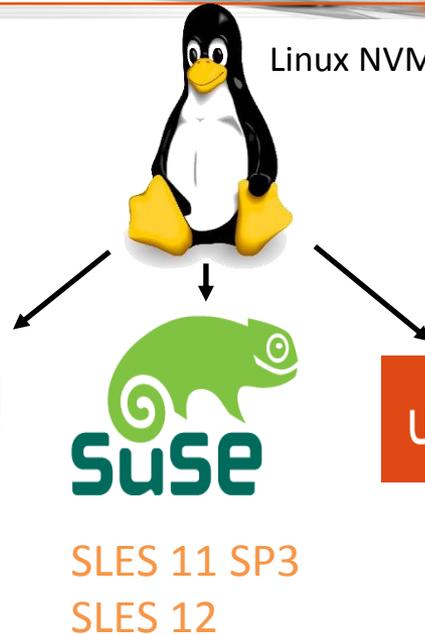
Native / in-box



Windows 8.1



redhat
6.5 | 7.0



Linux NVMe driver is [open source](#)



SLES 11 SP3
SLES 12



13 | 14



Install NVMe driver



NVM Express Advantages over SATA



PCIe for **scalable** performance, **flexible** form factors, and industry **stability**



NVMe provides **lower latency** and increased **efficiency**: lower CPU utilization, lower power, lower TCO



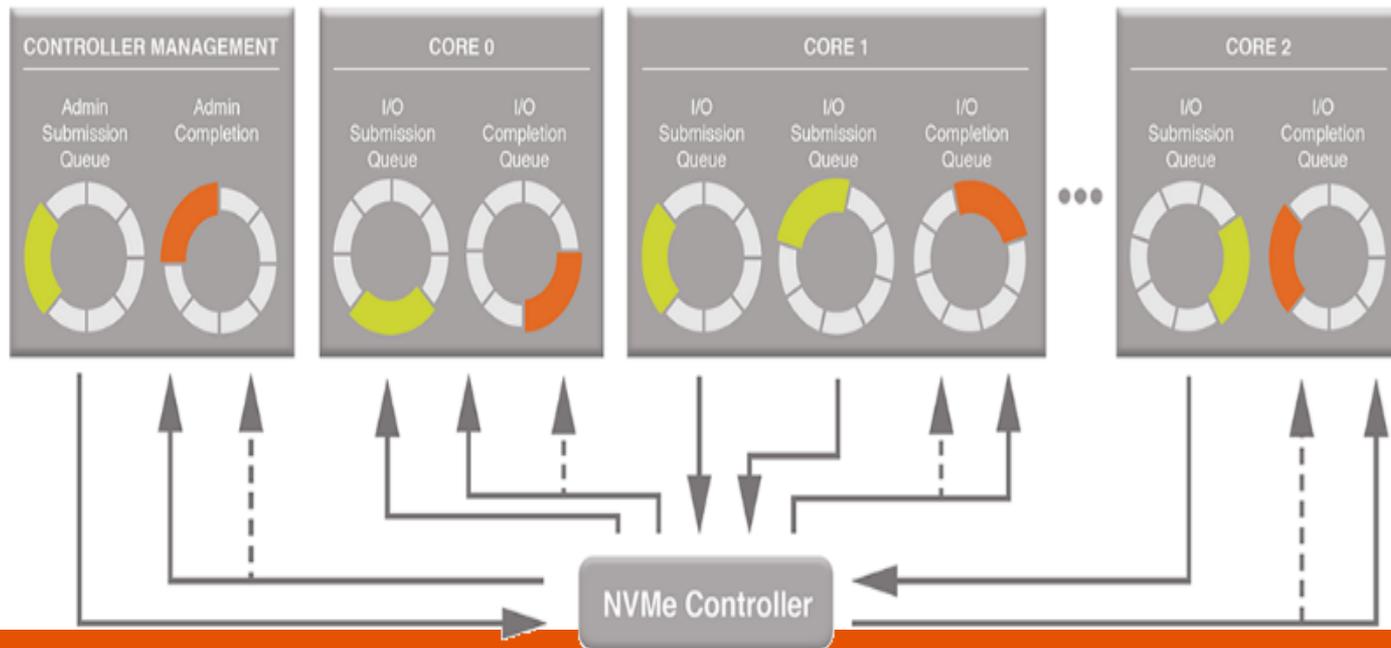
Increased **bandwidth**: 1 GB/s per lane – 1-16 lanes per drive
Directly attached to CPU, eliminate HBA cost and overhead



Low power features from both PCIe and NVMe
Security from Trusted Computing Group OPAL

NVM Express Technical Overview

- Supports deep queues of 64K commands per queue, up to 64K queues
- Supports MSI-X and interrupt steering, enables even performance scaling
- Streamlined & simple command set (13 required commands), optional features to address target segments
- Built for the future, ready for next gen NVM



NVM Express (NVMe) Delivers Best in Class IOPs ... And Best in Class Sequential Performance



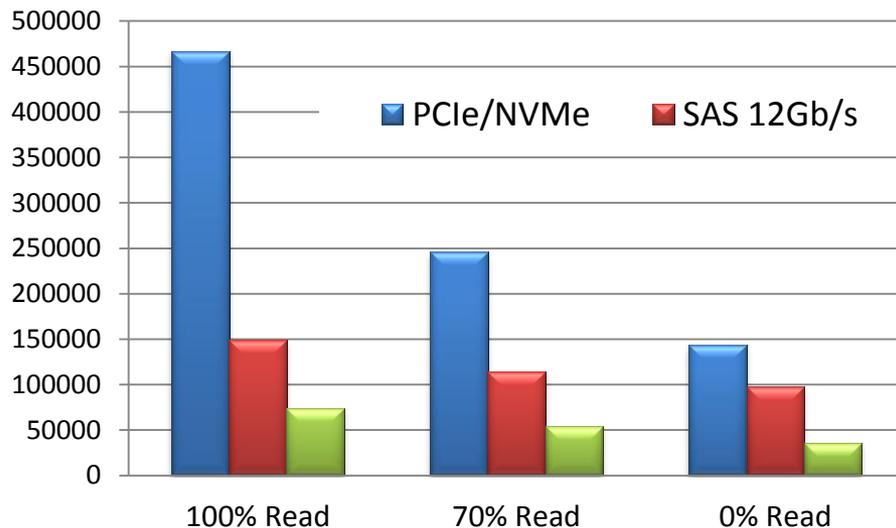
Compared to SAS 12 Gbps

- 100% random reads: >3X better IOPs
- 70% random reads: >2X better IOPs
- 100% random writes: ~1.5X better IOPs

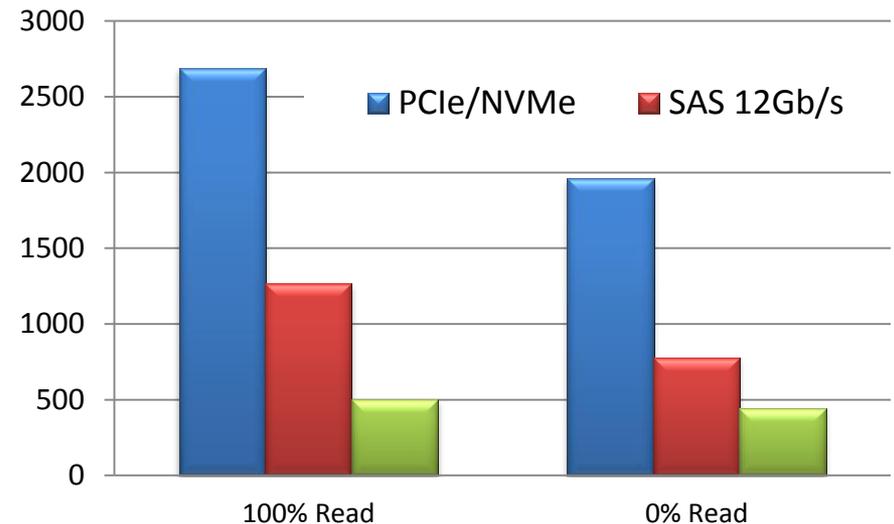
Compared to SAS 12 Gbps

- 100% reads: >2X better performance
- 100% writes: >2.5X better performance

4K Random Workloads (IOPs)



Sequential Workloads (MBps)



Note: PCI Express® (PCIe®)/NVM Express® (NVMe) Measurements made on Intel® Core™ i7-3770S system @ 3.1GHz and 4GB Mem running Windows® Server 2012 Standard O/S, Intel PCIe/NVMe SSDs, data collected by Iometer® tool. PCIe/NVMe SSD is under development. SAS Measurements from HGST Ultrastar® SSD800M/1000M (SAS) Solid State Drive Specification. SATA Measurements from Intel Solid State Drive DC P3700 Series Product Specification.

Analyzing What Matters

- What matters in today's Data Center is not just IOPs and bandwidth
- Let's look at efficiency of the software stack, latency, and consistency

Server Setup



- Basic 4U Intel® Xeon® E5 processor based server
- Out of box software setup
- Moderate workload: 8 workers, QD=4, random reads

Storage Protocols Evaluated

Interface	6Gb SATA	6Gb SATA	6Gb SAS	12Gb SAS	NVMe PCIe Gen 3
Attach Point	PCH chipset	6Gb SAS HBA	6Gb SAS HBA	12Gb SAS HBA	CPU

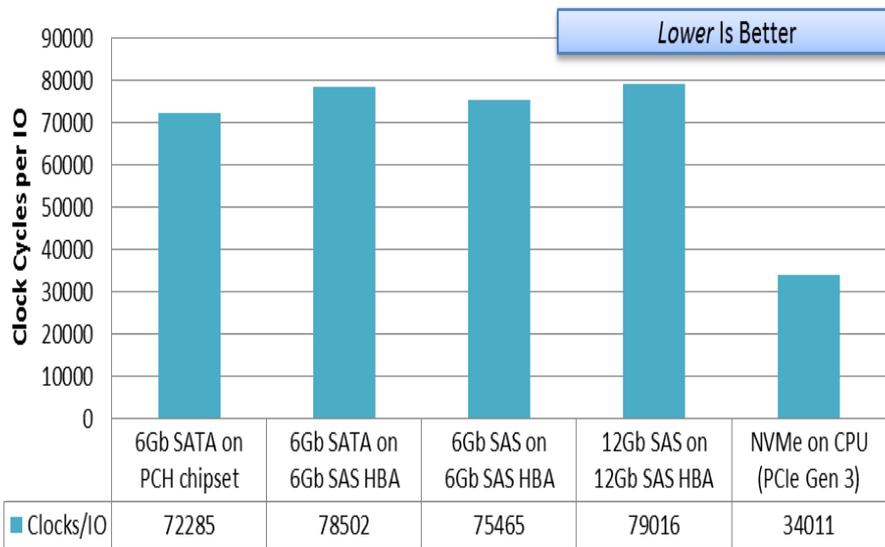
NVM Express* (NVMe)
PCI Express* (PCIe*)

Not strenuous on purpose – evaluate protocol and not the server.

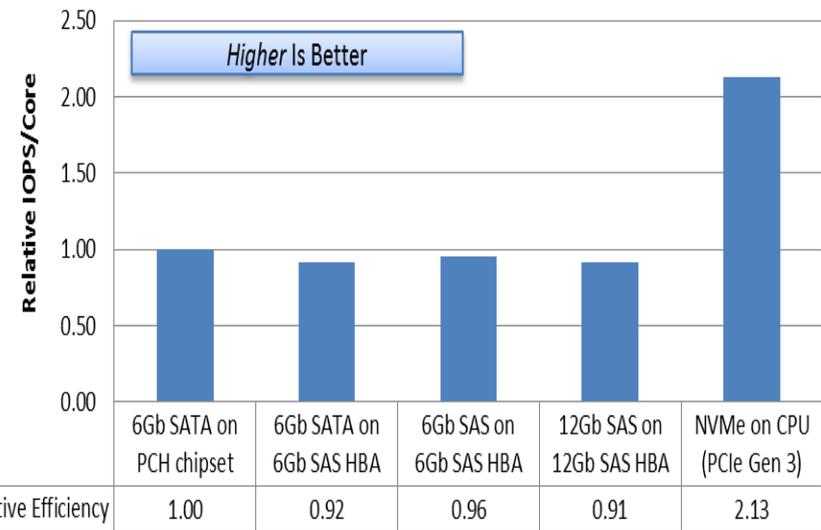
The Efficiency of NVM Express™

- CPU cycles in High Performance Computing are precious
 - Each CPU cycle required for an IO adds latency
- NVM Express takes less than half the CPU cycles per IO as SAS

CPU Cycles per IO



Relative Efficiency

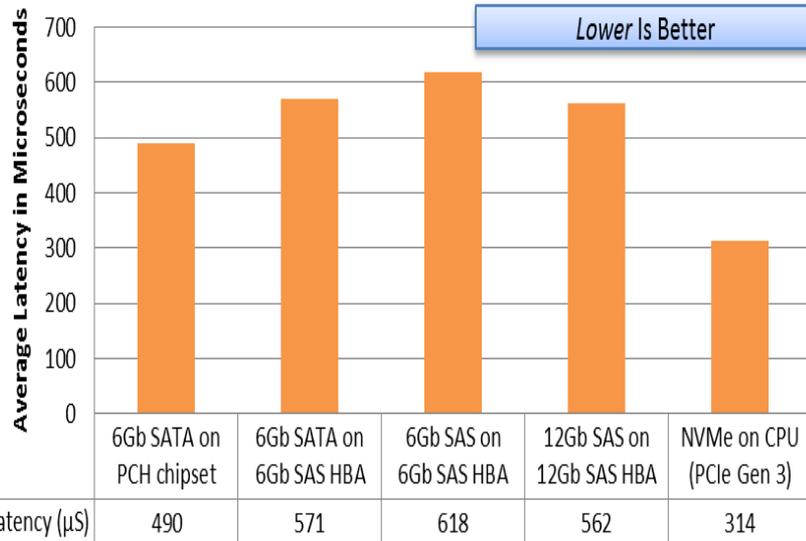


With equivalent CPU cycles, NVM Express delivers over 2X the IOPs of SAS!

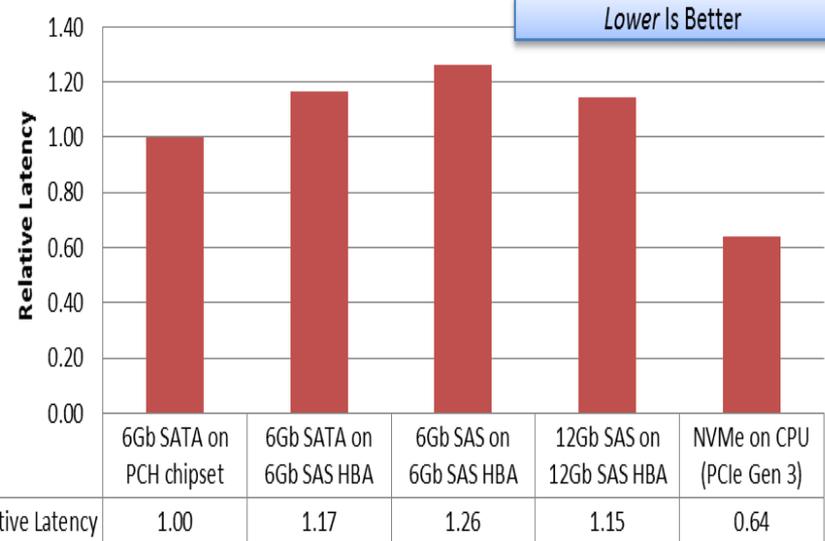
The Latency of NVM Express™

- The efficiency of NVM Express directly results in leadership latency
- When doubling from 6Gb to 12Gb, SAS only reduces latency by ~ 60 μ S
- NVMe is more than 200 μ s lower average latency than 12 Gb SAS

Average Latency in μ S



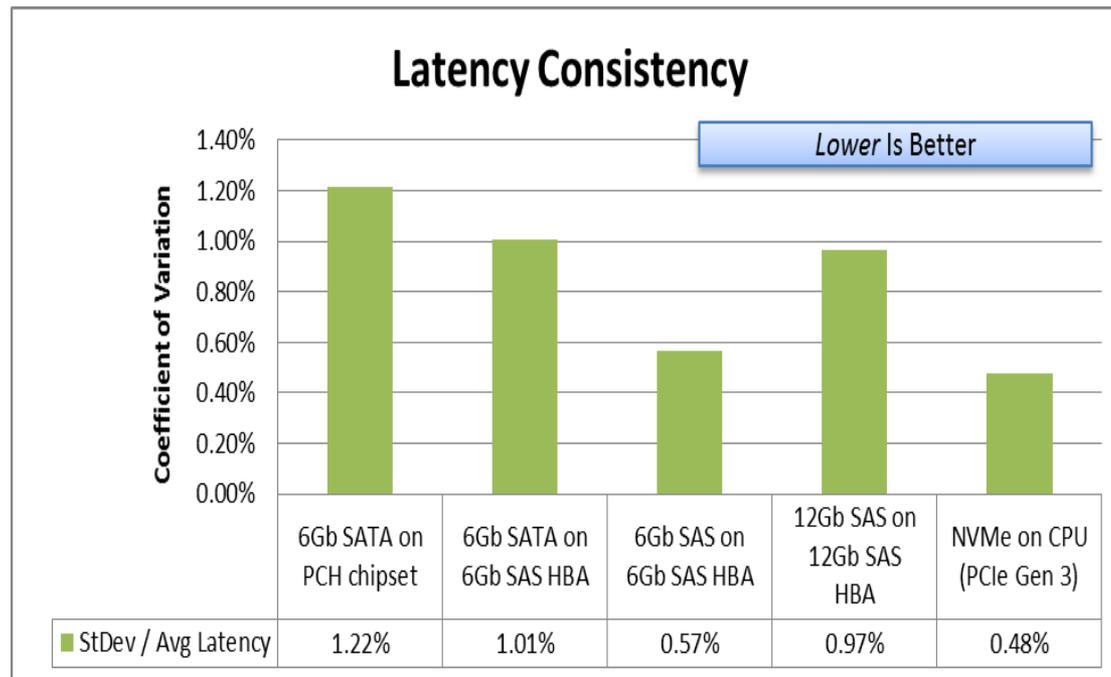
Relative Latency



NVM Express delivers the lowest latency of any standard storage interface.

The Consistency of NVM Express*

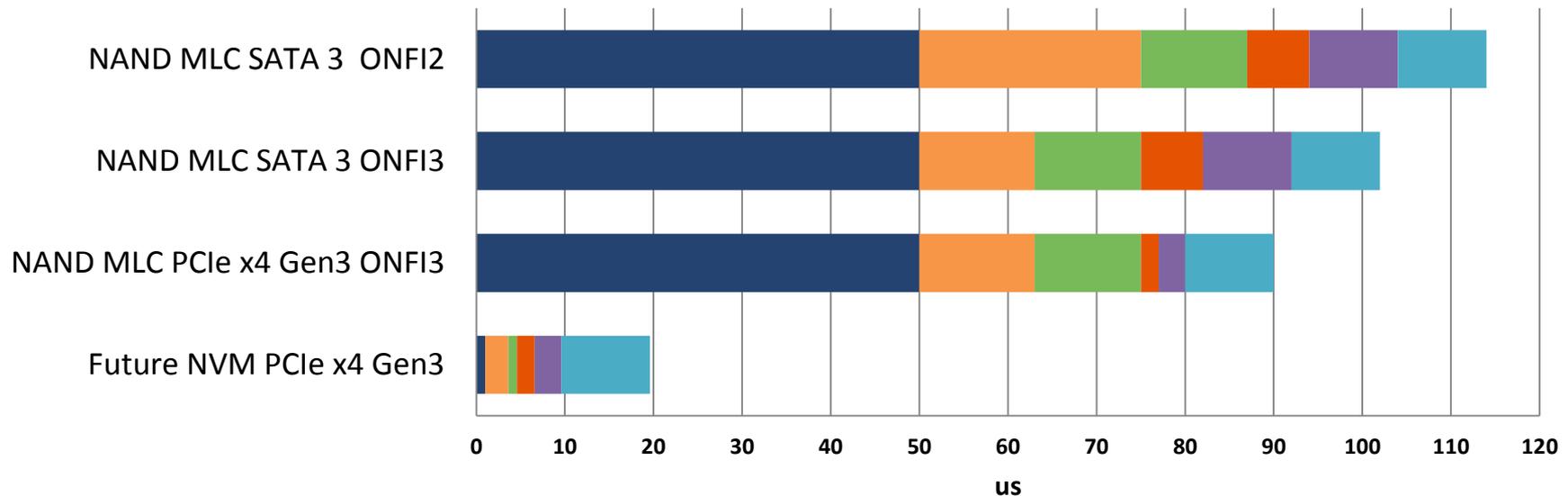
- NVM Express* (NVMe) leadership on latency and efficiency is **consistently** amazing
- SAS is a mature software stack with over a decade of tuning, yet the first generation NVM Express software stack has 2 to 3X better consistency



NVMe is already best in class, with more tuning yet to come.

Fully Exploiting Next Gen NVM

- With Next Gen NVM, the NVM is no longer the bottleneck
- App to SSD read latency for 4KB transfer at Queue Depth of 1



The choice is yours...



OPENFABRICS
ALLIANCE



For more information...



The screenshot shows the top section of the NVM Express website. At the top left is the NVM Express logo. Below it is a navigation bar with icons and labels for HOME, ABOUT, NEWS, PRODUCTS, RESOURCES, MEMBERSHIP, and BLOG. The main banner features a colorful, abstract background with the text "NVM EXPRESS" in large white letters. A small "TM" trademark symbol is visible to the right of the word "EXPRESS".

The Optimized PCI Express® SSD Interface

The NVM Express specification defines an optimized register interface, command set and feature set for PCI Express (PCIe®)-based Solid-State Drives (SSDs). The goal of NVM Express is to unlock the potential of PCIe SSDs now and in the future, and standardize the PCIe SSD interface.

Visit nvmexpress.org



Thank You



#OFADevWorkshop