SNIA SSSI | SOLID STATE STORAGE

NVDIMM - the Go-To Technology for Boosting System Performance

Arthur Sainio
SNIA NVDIMM Special Interest Group
SMART Modular Technologies

August 8, 2016

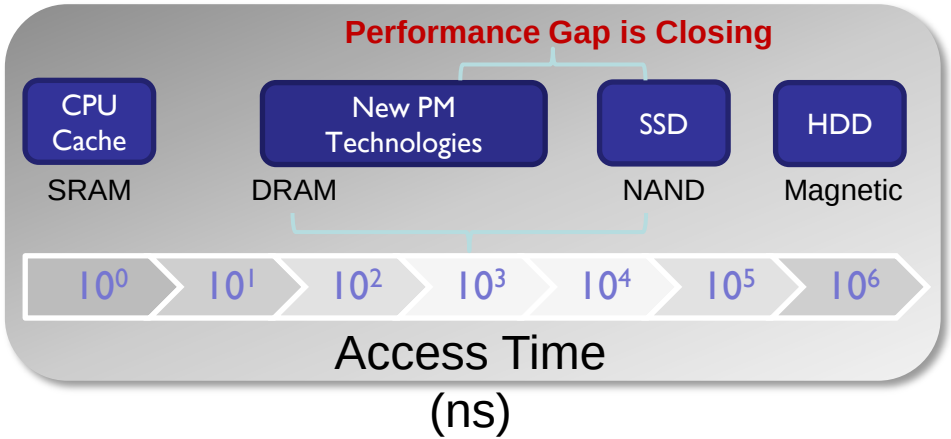
NVDIMM – The Go-To Technology Topics

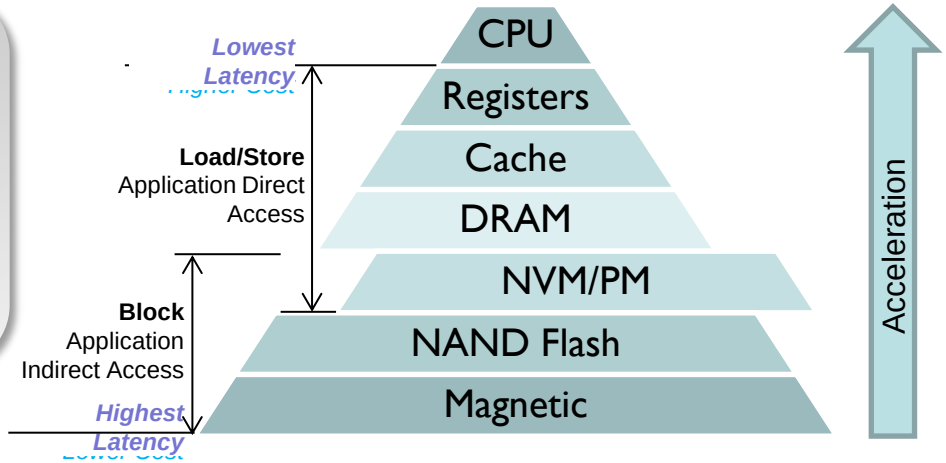
SNIA SSSI | SOLID STATE STORAGE

- How are NVDIMMs a revolutionary technology which will boost the performance of next-generation server and storage platforms?
- What are the standardization and ecosystem enablement efforts around NVDIMMs that are paving the way for plug-n-play adoption?
- What would customers, storage developers, and the industry like to see to fully unlock the potential of NVDIMMs?

Memory – Storage Hierarchy

- Data-intensive applications need fast access to storage
- Persistent memory is the ultimate high-performance storage tier
- NVDIMMs have emerged as a practical next-step for boosting performance

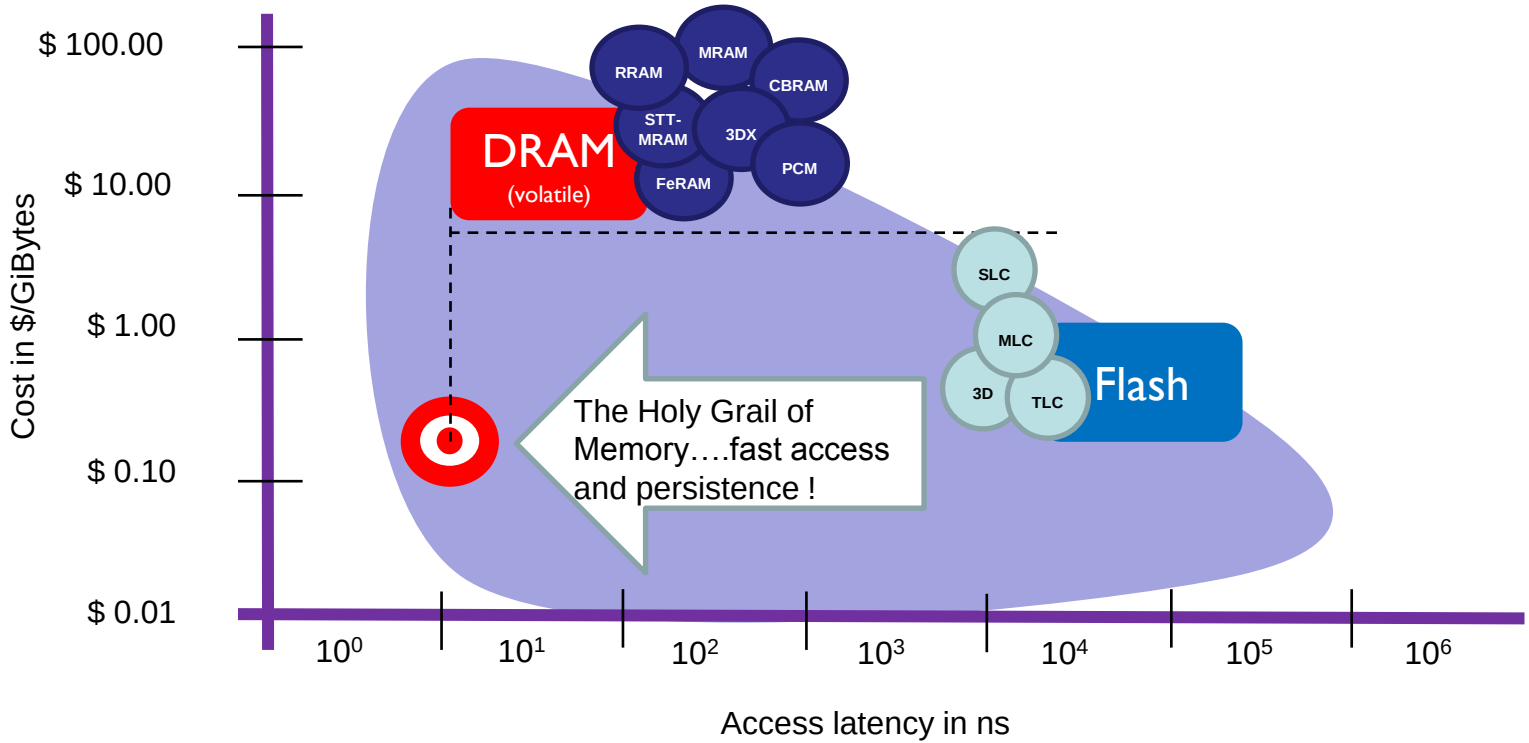




Source; HPE/SNIA 2015

Persistent Memory Types

Room for multiple Types



Source; Adapted form HPE/SNIA 2016

Application Opportunities with Persistent Memory

- Performance
  - Lighter software stacks
  - Direct memory access
  - Better CPU utilization
  - Cache acceleration
- Capacity
  - Transaction logging
  - Hot data sets for analytics, in-memory computing
- Endurance
  - Realize performance and persistence values for a wide range for work loads
- Persistence
  - Converge storage and memory

Source; Adapted form Parallel Machines, Storage Visions Conference, Jan'16

NVDIMMs - JEDEC Taxonomy

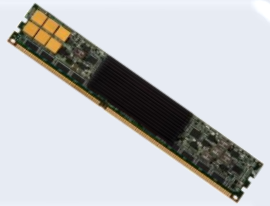
SNIA SSSI | SOLID STATE STORAGE

NVDIMM-N
Standardized

- Memory mapped DRAM. Flash is not system mapped
- Access Methods -> byte- or block-oriented access to DRAM
- Capacity = DRAM DIMM (1's -10's GB)
- Latency = DRAM (10's of nanoseconds)
- Energy source for backup
- DIMM interface (HW & SW) defined by JEDEC

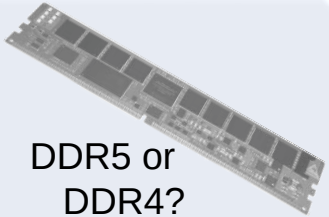
NVDIMM-F
Vendor Specific

- Memory mapped Flash. DRAM is not system mapped.
- Access Method -> block-oriented access to NAND through a shared command buffer (i.e. a mounted drive)
- Capacity = NAND (100's GB-1's TB)
- Latency = NAND (10's of microseconds)



NVDIMM-P
Proposals in progress

- Memory-mapped Flash and memory-mapped DRAM
- Two access mechanisms: persistent DRAM (–N) and block-oriented drive access (–F)
- Capacity = NVM (100's GB-1's TB)
- Latency = NVM (100's of nanoseconds)



DDR5 or DDR4?

NVDIMM-N Combines the Best of Flash & DRAM

HDD

IOPS
(even if random…)

FLASH

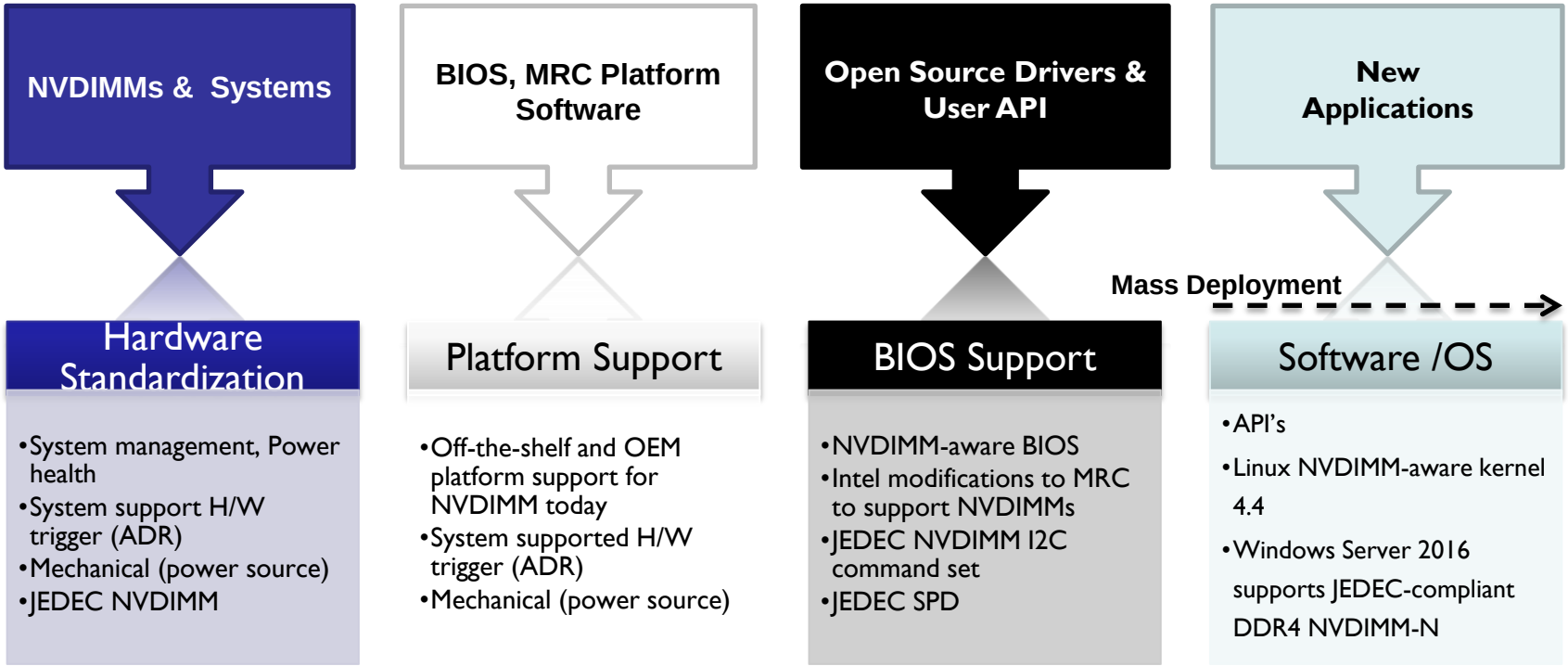
Latency
(even under load…)

NVDIMM

DRAM Memory Speed + **NAND Storage Persistency** = **Lower Latency Higher IOPS = Improved System Performance**

- Many NVDIMM-N enabled systems are available and shipping now
- Many NVDIMM-N vendors are providing support

NVDIMM-N Ecosystem



NVDIMM-N Development Updates

- JEDEC DDR4 Standardization
  - SAVE_n: pin 230 sets a efficient interface to signal a backup
  - 12V: pin 1, 145 provides power for backup energy source
  - EVENT_n: pin 78 asynchronous event notification pin
  - Byte Addressable I2C interface (JESD245)
  - JEDEC defined SPD/Registers to comply with DDR4 RDIMM
- NVDIMM firmware interface table (NFIT) added in ACPI 6.0 (Advanced Configuration and Power Interface Specification)
- Intel MRC/BIOS supports JEDEC I2C command set

NVDIMM-F Motivation/Challenges Block Accessed NAND Flash

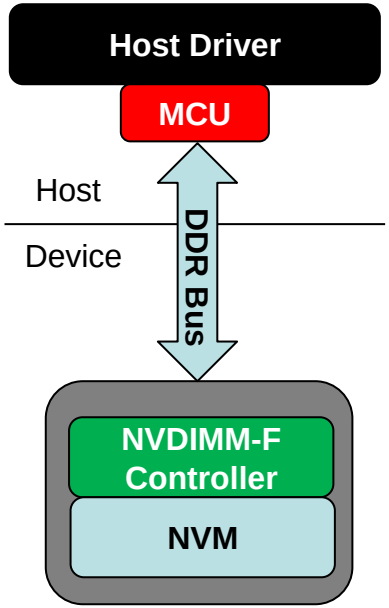
SNIA SSSI | SOLID STATE STORAGE

Motivation

- Moving NAND to memory channel eliminates traditional HDD/SSD SAS/PCIe link transfer, driver, and software overhead. As storage latency decreases these factors dwarf the storage access percentage of an read/write.
- DDR interface directly to NVM
- Enables hundreds of GBs per DIMM
- Enables tens of TBs per server
- Leverages economic advantages of NVM within memory subsystem

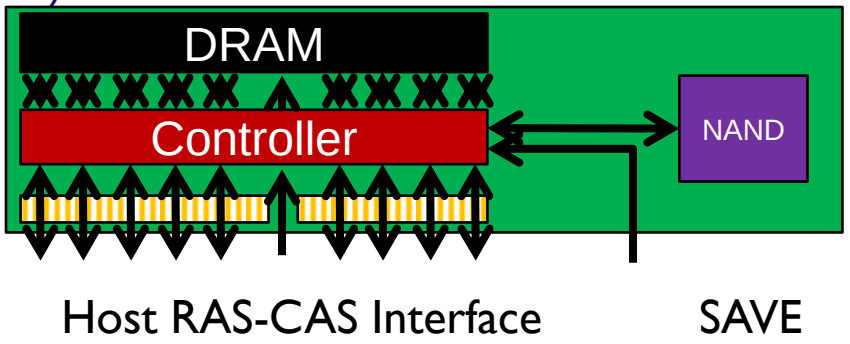
Challenges

- NAND 10,000x slower than DRAM. Attachment to memory channel must not interfere with DRAM performance
- NAND block access vs. DRAM byte access



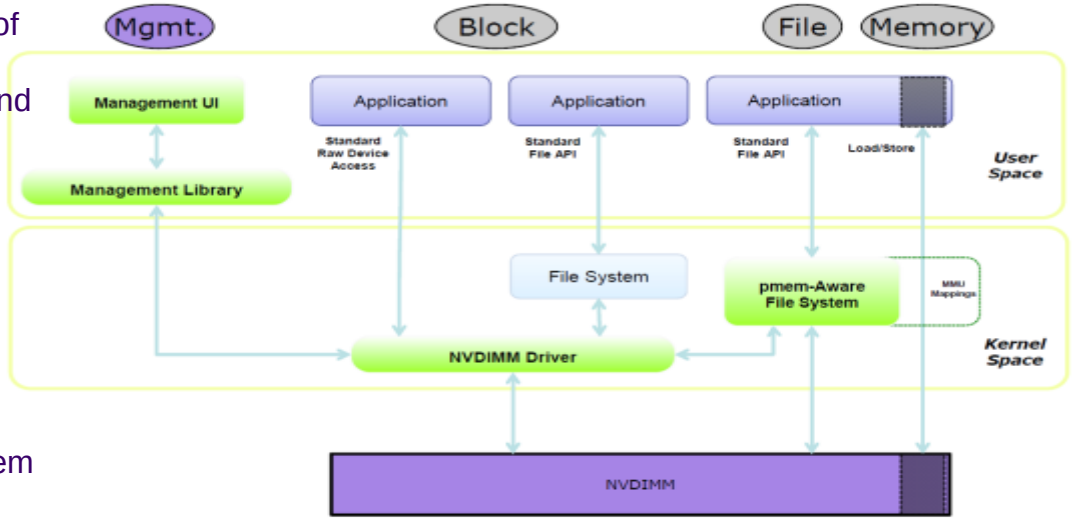
NVDIMM-P Combines DRAM & Flash

- Memory-mapped Flash and memory-mapped DRAM
- Two access mechanisms: persistent DRAM (–N) and block-oriented drive access (–F)
- Capacity - 100's GB to 1's TB
- Latency 100's of nanoseconds
- NVDIMM-P definition in discussion
- Existing DDR4 protocol supported
- Extensions to protocol under consideration
  - Sideband signals for transaction ID bus
  - Extended address for large linear addresses



SNIA NVM Programming Model

- Developed to address the ongoing proliferation of new NVM technologies
- Necessary to enable an industry wide community of NVM producers and consumers to move forward together through a number of significant storage and memory system architecture changes
- The specification defines recommended behavior between various user space and operating system (OS) kernel components supporting NVM
- The specification does not describe a specific API. Instead, the intent is to enable common NVM behavior to be exposed by multiple operating system specific interfaces



Source; Parallel Machines, Storage Visions Conference, Jan'16

Application Access to NVDIMMs

- Block Storage
- Disk-like NVDIMMs (-F or -P)
- Appear as disk drives to applications
- Accessed using disk stack
- Block Mode
  - Low latency
  - Compatible with existing file system and storage drivers

- Direct Access Storage (DAS)
- Memory-like NVDIMMs (-N or -P)
- Appear as memory to applications
- Applications store variables directly in RAM
- No IO or even DMA is required
- Absolute lowest latency (fastest server performance)
- No OS between the application and the SCM
- Byte addressable storage

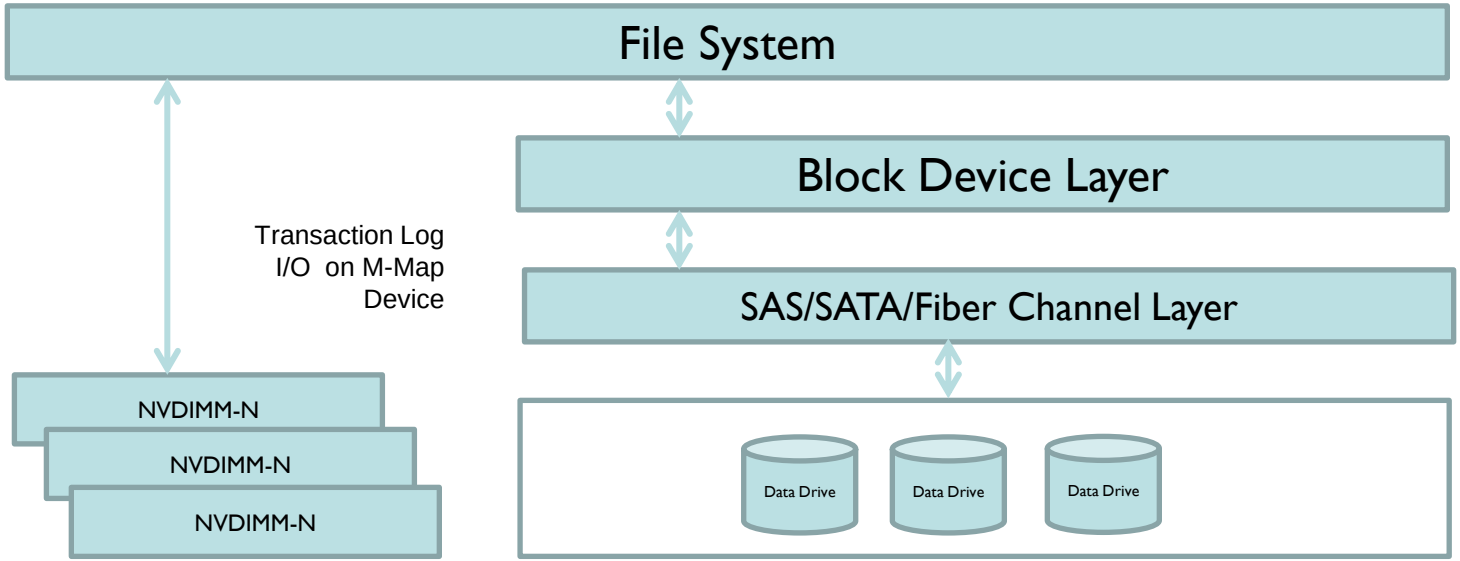
Applications Enabled by the NVM Programming Model

SNIA SSSI | SOLID STATE STORAGE

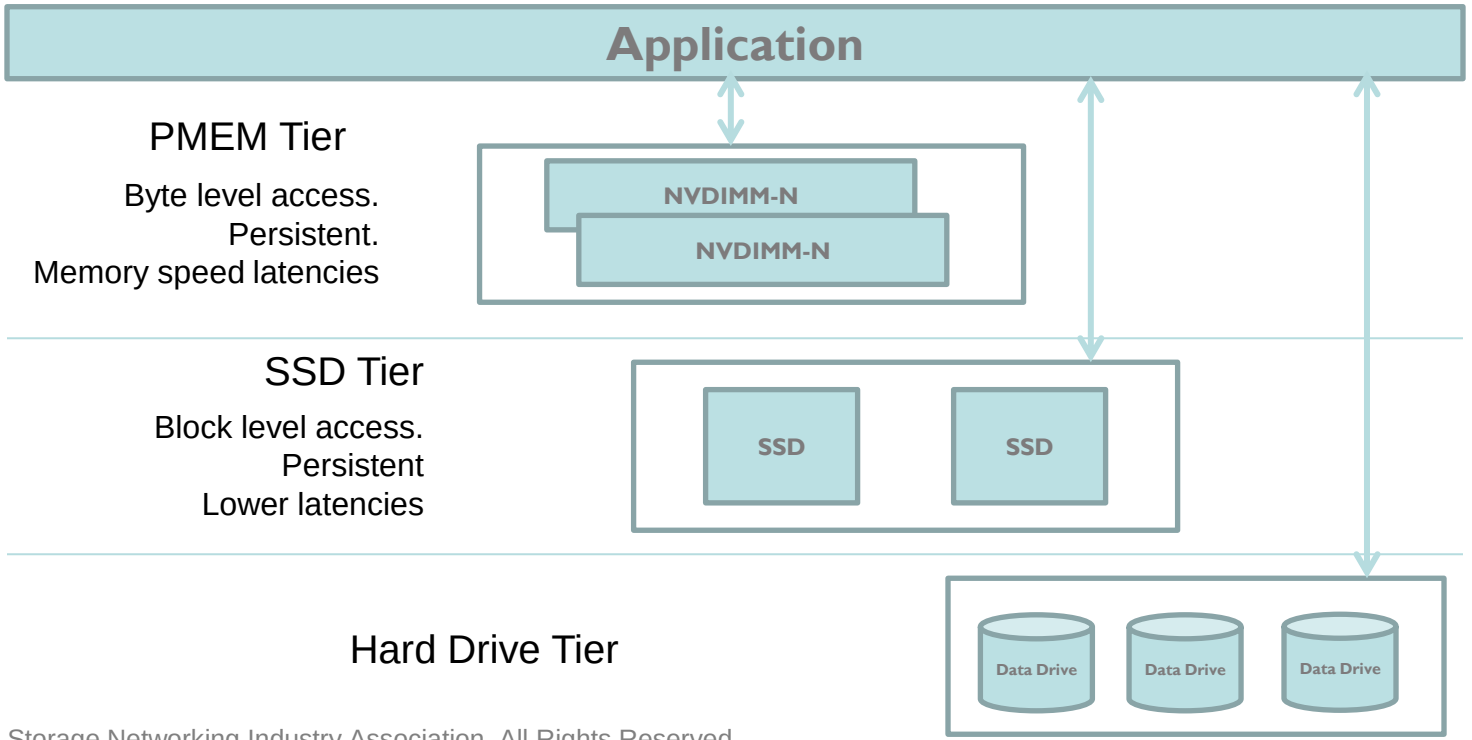
- File Systems
  - Metadata/log acceleration, data tiering, whole persistent memory file systems
- Databases and In-Memory
  - Small capacity – caching, log acceleration
  - Larger capacity – drive larger transaction rates, in-memory databases with persistence
- Analytics and Machine Learning
  - Larger dataset sizes, greater information processing, improved machine learning accuracy
  - Converge analytics and real time data processing

Source; Adapted from HPe, Storage Visions Conference, Jan'16

NVDIMM-N Use Case #1
File System Transaction Log



NVDIMM-N Use Case #2
Application Persistent Data Tier



Linux & Microsoft - NVDIMM-N

SNIA SSSI | SOLID STATE STORAGE

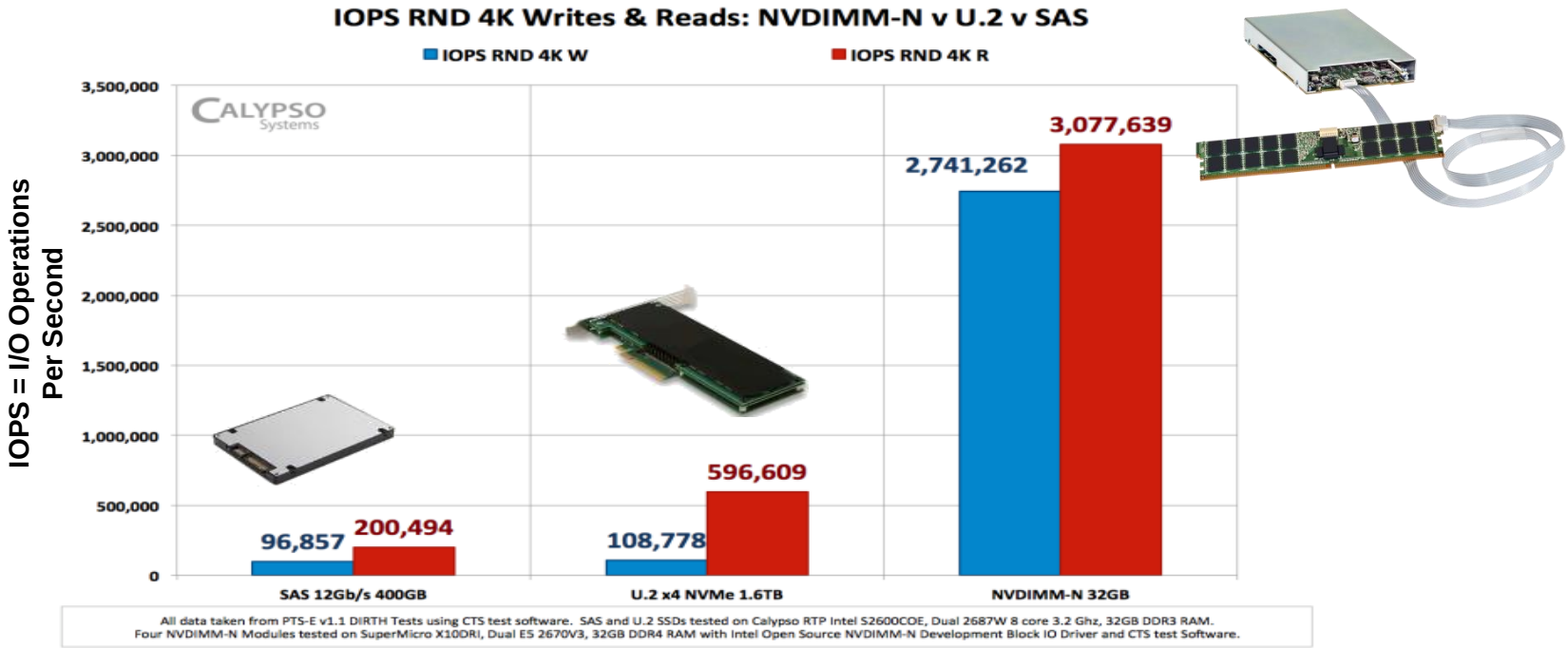
Linux

- Persistent Memory in Linux
- Linux 4.4 subsystems added and modified in support of NVDIMMs
- Core Kernel support for ACPI 6.0 with NFIT BIOS, Device Drivers, Architectural Code, and File System with DAX support (ext4)
- Distributions (Open Source Initiatives)
  - Ubuntu 16.04 LTS (4.4 Kernel)
  - Fedora 23 (4.2.0 Kernel)

Microsoft

- At this year's //Build conference MS made public that Windows Server 2016 supports JEDEC-compliant DDR4 NVDIMM-N
  - https://channel9.msdn.com/Events/Speakers/tobias-klima
- Technical Preview 5 of Windows Server 2016, has NVDIMM-N support
  - https://www.microsoft.com/en-us/evalcenter/evaluate-windows-server-technical-preview)

NVDIMM-N Benchmark Testing



Source: Calypso Systems

NVDIMM-N Benchmark Testing and Multivendor NVDIMM Demo

SNIA SSSI | SOLID STATE STORAGE

- Showing performance benchmark testing using a SDM (Software Defined Memory) file system
- Compares the performance between four 16GB DDR4 NVDIMMs and a 400GB NVMe PCIe SSD
- The NVDIMMs create a byte-addressable section of persistent memory within main memory allowing for high-speed DRAM access to business-critical data
- Demo
  - Motherboard - Supermicro X10DRi
  - Intel E5-2650 V3 processor
  - Four 16GB NVDIMMs and supercap modules (Micron, Netlist, SMART)
  - Four 16GB RDIMMs
  - One Intel 750 series 400GB NVMe PCIe SSD
  - Plexistor SDM file system

What Customers, Storage Developers, and the Industry Would Like to See to Fully Unlock the Potential of NVDIMMs

- **Standardization and Interoperability**
  - Standard server and storage motherboards enabled to support all NVDIMM types
  - Standardized BIOS/MRC, driver, and library support
  - Interoperability between MBs and NVDIMMs
  - Standardized memory channel access protocol adopted by Memory Controller implementations
  - O/S recognition of APCI 6.0 (NFIT) to ease end user application development
- **Features**
  - Data encryption/decryption with password locking JEDEC standard
  - Standardized set of OEM automation diagnostic tools
  - NVDIMM-N Snapshot: JEDEC support of NMI trigger method alternative to ADR trigger
- **Performance**
  - Standardized benchmarking and results
  - Lower latency I/O access < 5us

SNIA SSSI | SOLID STATE STORAGE

Questions?
Thank you!