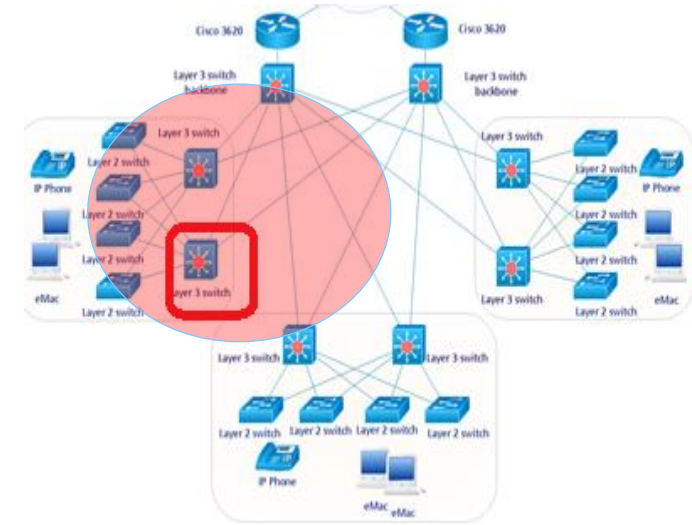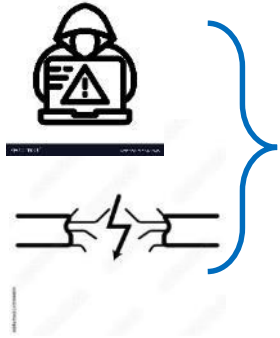# Scaling NVDIMM-N Architecture for System Acceleration in DDR5 and CXL™ - Enabled Systems

Presented by; Arthur Sainio and Pekon Gupta

Need for Persistent Memory

Persistence

+

Memory (low latency)

- System downtimes are costly but inevitable.

- To limit the blast radius, impacted services, VM or a downed server should be rebooted in the same state as it was earlier.

- Intermediate transaction states, configurations and metadata are saved in Persistent Memory even before transactions are fully committed.

- To identify the cause of failure, error logging until the latest event and security logging during a malicious hack are required.

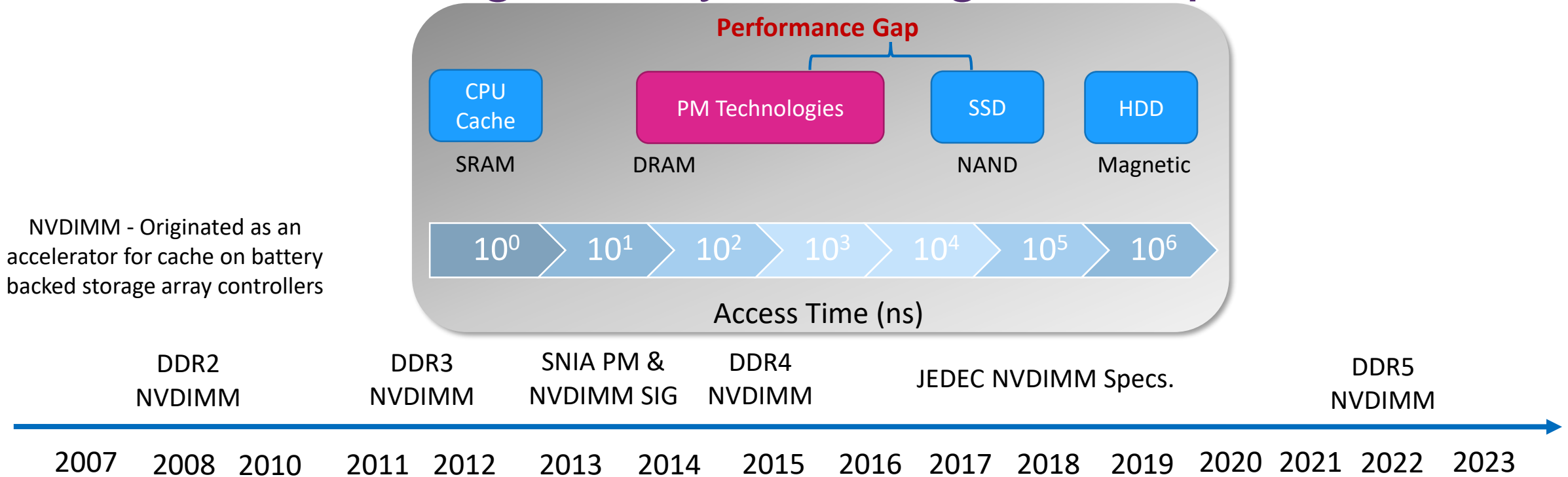- Persistence preserves event logs across power-cycles.

- High latency may directly impact revenue or customer experience.

- Missed SLA like denial of transaction in credit card business means losing the transaction to your competitor or allowing fraudulent transaction.

- In High Frequency trading, A broker could lose millions, if their electronic trading platform is just few milliseconds behind the competition.

- Amazon found every 100ms of latency cost them 1% in sales.

- Google found an extra 0.5 seconds in search page generation time dropped traffic by 20%.

PERSISTENT MEMORY
SNIA + SUMMIT 2022
COMPUTATIONAL STORAGE

Source; The GigaSpaces Technologies Blog, Insights into In-Memory Computing and Real-time Analytics

Persistent Memory Solution Options

|	Volatile Memory with Backup Power (NVDIMM-N)	Non-Volatile Media Based (NVDIMM-P)
Architecture		 • PCM: Phase Change Memory
Examples	• NVDIMM-N • CXL™ NV-XMM implementations (Non-Volatile CXL Memory Modules)	• Intel Optane™ PMem • NVDIMM-P based implementations • Different types of PM media
Features	✓ Low latency (RTT < 50ns) (Incoming traffic hits DRAM) ✓ Granular Byte level addressing. ✓ High reliability (Flash used only during backup) ✓ Data retention for data at rest equivalent to NAND media. X Hold up power required	X Higher Latency ✓ Granular Byte level addressing X Reliability (Limited by endurance of media) ✓ Data retention across power-cycles ✓ High density memory expansion

PERSISTENT MEMORY + SUMMIT 2022
SNIA COMPUTATIONAL STORAGE

NVDIMMs – Long History of Filling the Gap

Performance Gap

CPU Cache	PM Technologies	SSD	HDD
SRAM	DRAM	NAND	Magnetic

10^0 > 10^1 > 10^2 > 10^3 > 10^4 > 10^5 > 10^6

Access Time (ns)

NVDIMM - Originated as an accelerator for cache on battery backed storage array controllers

DDR2 NVDIMM	DDR3 NVDIMM	SNIA PM & NVDIMM SIG	DDR4 NVDIMM	JEDEC NVDIMM Specs.	DDR5 NVDIMM

2007 2008 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023

- **Practical industry solution to dramatically increase system performance**
 - ✓ Provides direct access and removes IO and all the overhead
  - ✓ Allows critical databases to be built in memory
  - ✓ A memcached structure is faster than the any solid-state solution
  - ✓ Uses off-the-shelf DRAM and Flash components
  - ✓ NVDIMM solutions are widely adopted by flash storage array vendors.

SNIA PERSISTENT MEMORY + SUMMIT 2022 COMPUTATIONAL STORAGE

Use Cases - NVDIMMs in Storage
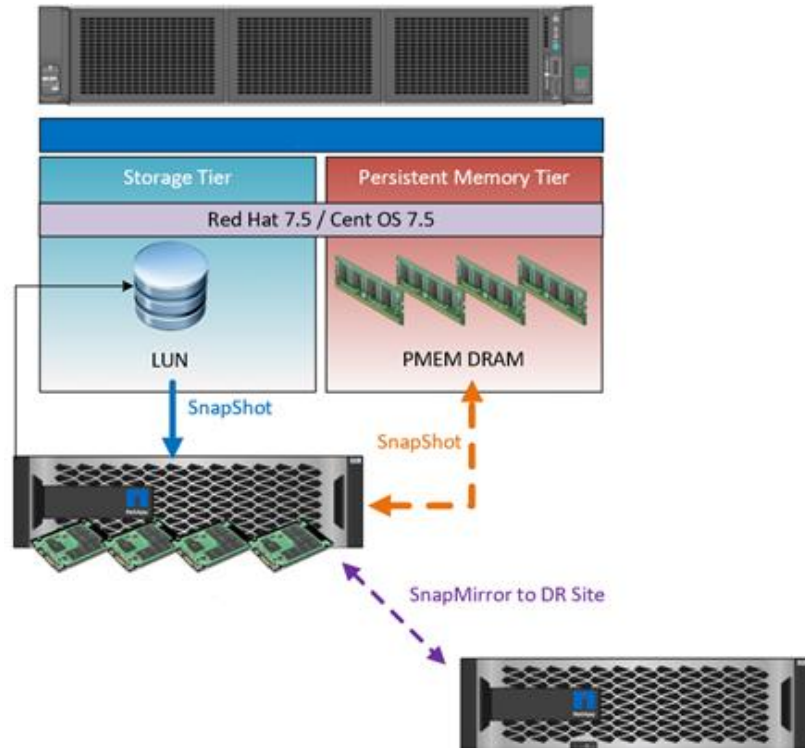
As read cache

- The NVDIMM solution is low in complexity from a software and hardware implementation perspective.
- NVDIMMs cache hot data and metadata, thereby enabling performance with latencies in the sub-80ns range, almost 100x faster less than high end SSD available.
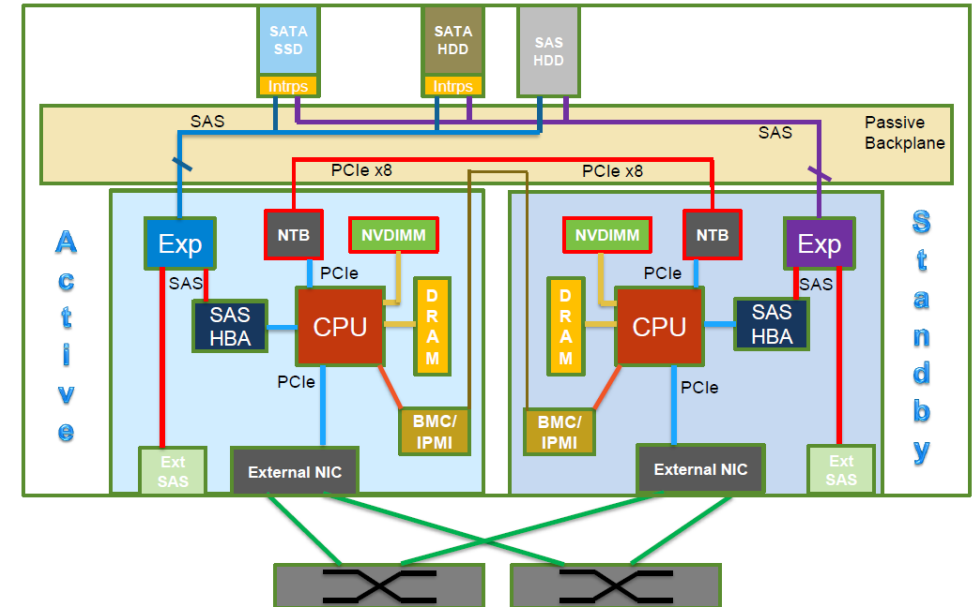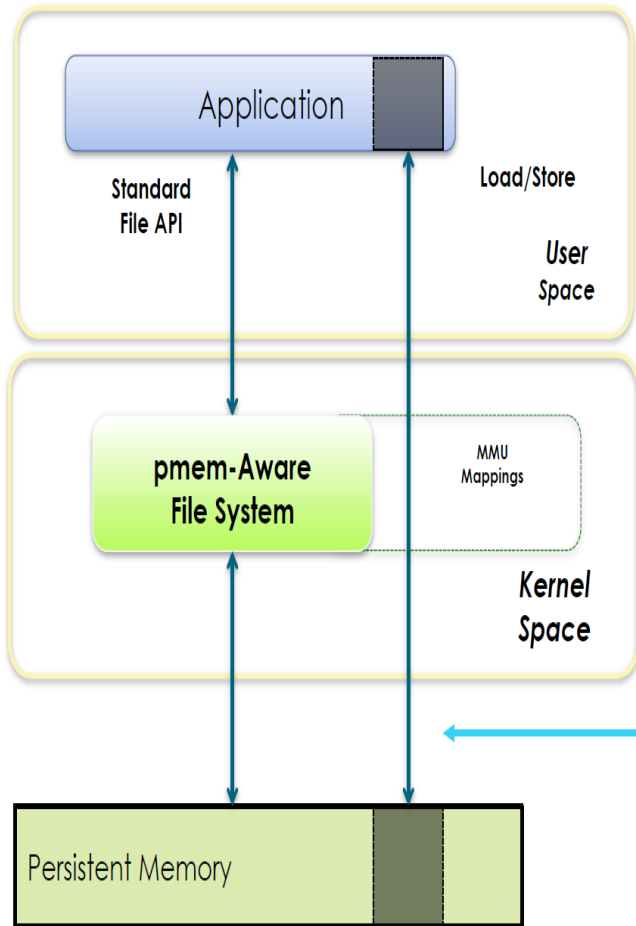

For Fault Tolerance

- Mirroring Data to and from NVDIMMs via PCIe NTB (Non-Transparent Bridging)
- NVDIMMs provide greater than 4X Write IOP latency improvement by reducing the time transient data must be replicated to the peer.

PERSISTENT MEMORY
SNIA + SUMMIT 2022
COMPUTATIONAL STORAGE

Established NVDIMM Ecosystem

SNIA Programming Model
Open source library for applications

Application

Standard File API

Load/Store

User Space

pmem-Aware File System

MMU Mappings

Kernel Space

DAX

Persistent Memory

- Linux 4.2 + added support of NVDIMMs.
- Windows Server 2016 supports JEDEC-compliant DDR4 NVDIMM-N

- NVDIMM-aware BIOS
- Intel modifications to MRC to support NVDIMMs
- JEDEC NVDIMM I2C command set
- JEDEC SPD

- Off-the-shelf and OEM platform support for NVDIMM today
- System supported H/W trigger (ADR)
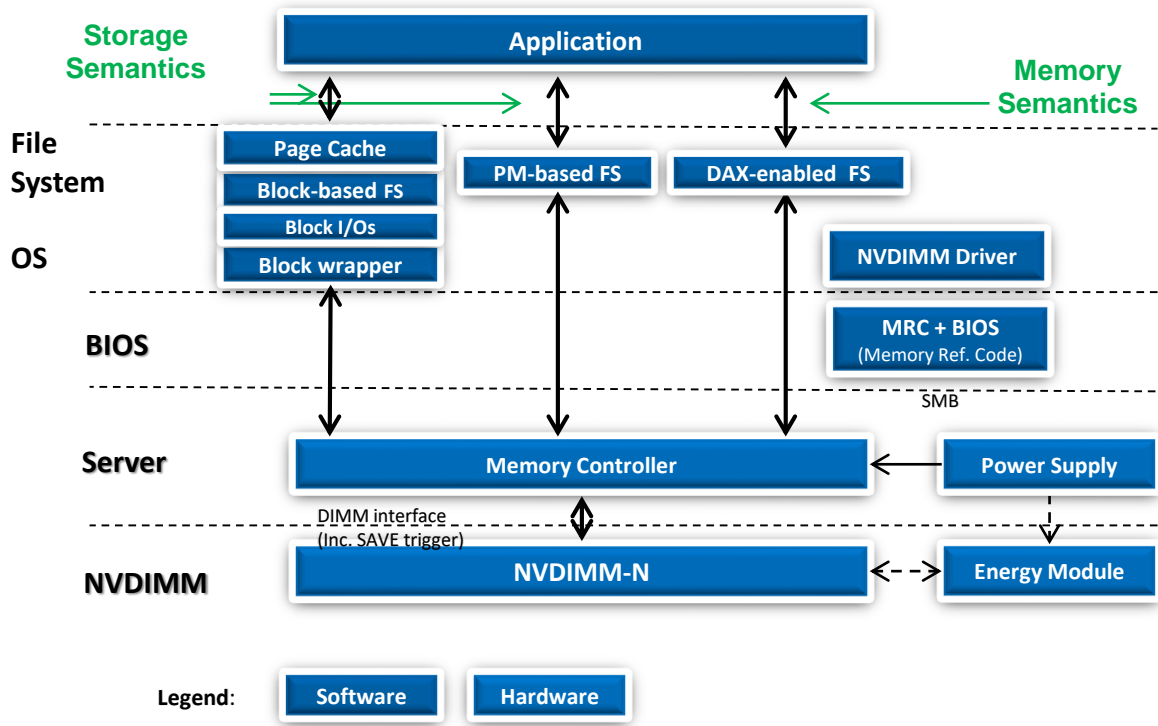- Mechanical (power source)

- System management, Power health Standardized through JEDEC and NFIT
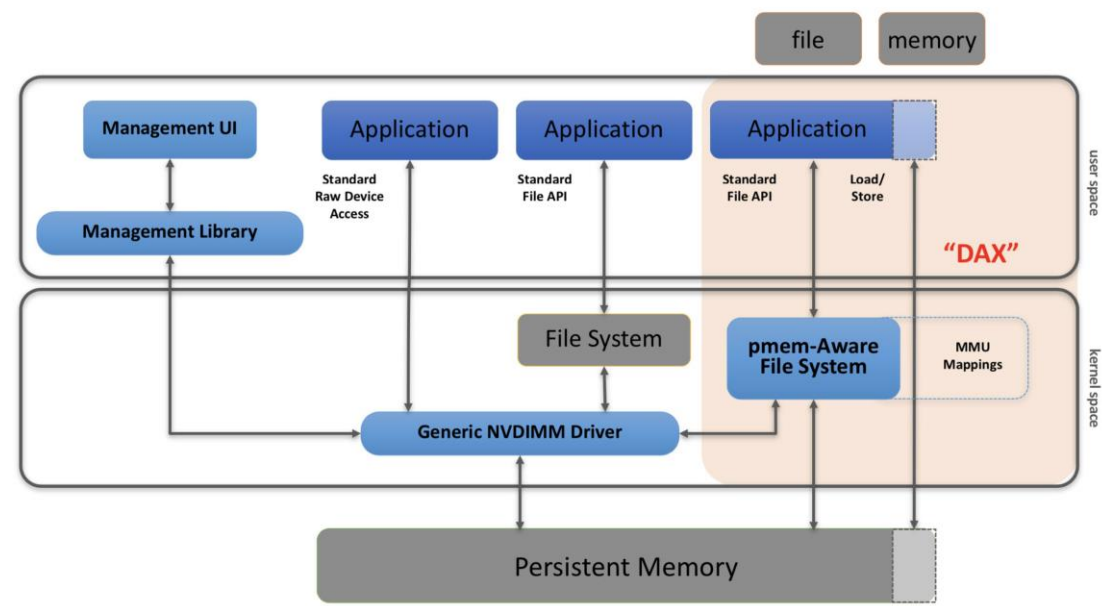- System support H/W trigger (ADR)
- Mechanical (power source)

Linux

UEFI
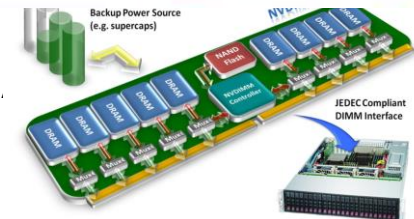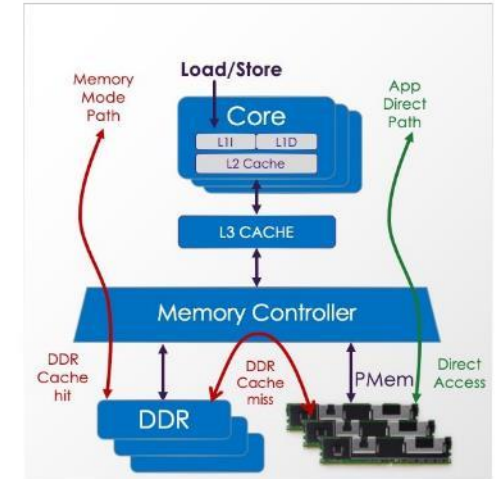
JEDEC®

JEDEC 245, 245B: Byte Addressable Energy Backed Interface

NVDIMM Cookbook – Still Valid for DDR5 NVDIMM

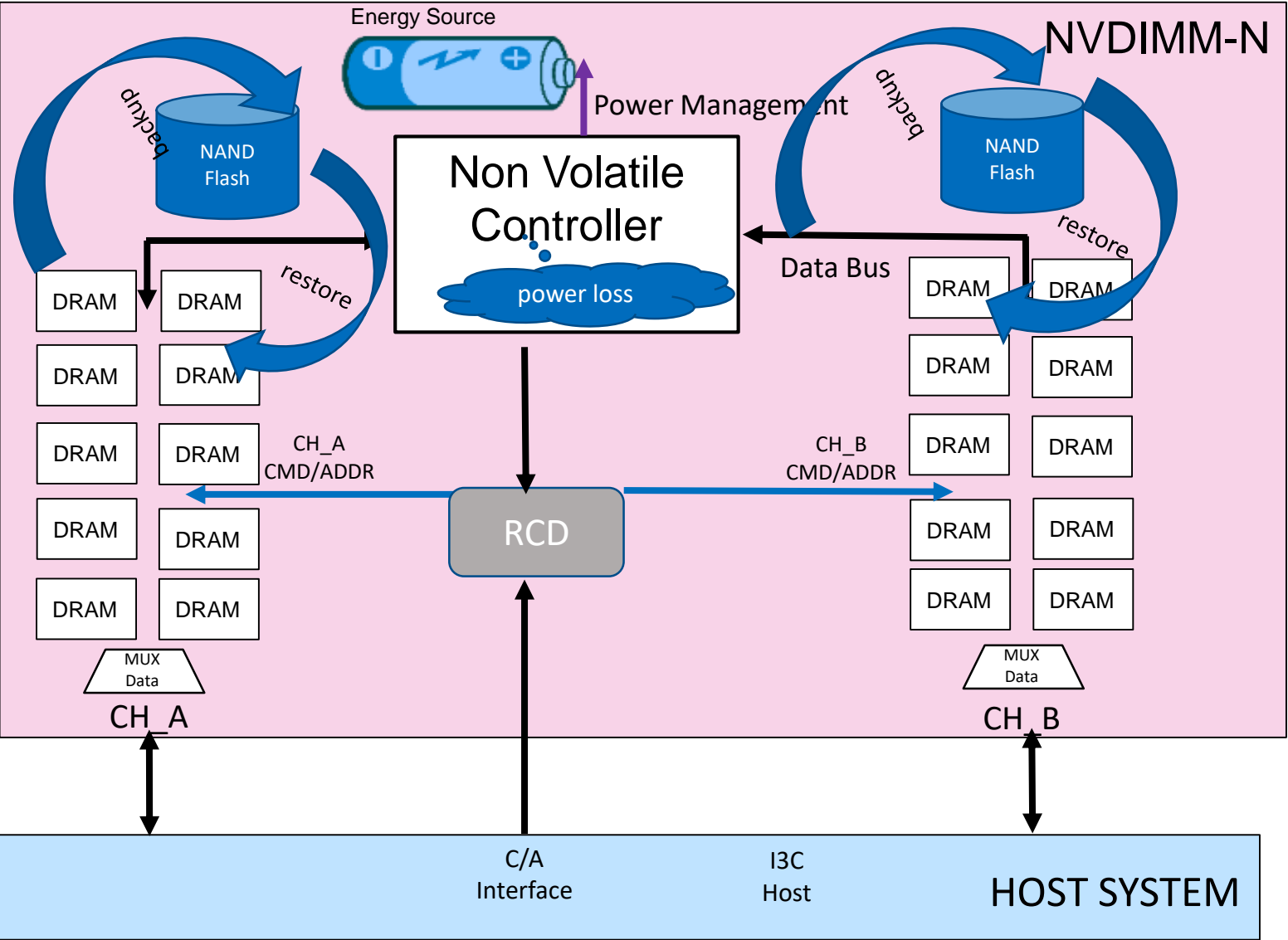The ACPI NVDIMM Framework
SNIA NVM Programming Model

NVDIMM Migration to DDR5

| Technology	DDR3	DDR4	DDR5	DDR5 Advantages
Speed	6.4 to 1.49 Mbps data rate	1.6 to 3.2Gbps data rate	4.8 to 6.4Gbps data rate	Higher Bandwidth. up to 6400 MHz
IO Voltage	1.5V	1.2V	1.1V	Lower Power
Power Management	On Mother board		On DIMM (PMIC)	Better power efficiency and scalability
Channel Architecture	72bit data channel (64 data + 8 ECC) 2 channels per DIMM	72bit data channel (64 data + 8 ECC) 1 channel per DIMM	40bit data channel (32 data + 8 ECC) 2 channels per DIMM	Higher memory efficiency, lower latency
Burst Length	4	8	16	Higher memory efficiency
Max. die density	8Gb	16Gb	32Gb	Higher capacity NVDIMMs
NVDIMM Types	SODIMM, RDIMM	RDIMM	LRDIMM, RDIMM	Same form factor

SNIA PERSISTENT MEMORY + SUMMIT 2022 COMPUTATIONAL STORAGE

DDR5 NVDIMM Architecture

Feature	Description
Memory	• DDR5-4800 32GB / 64GB
Protocol	• JEDEC Compliant DDR5
Features	• Throughput of 63.0 GB/s • Latency ~20ns • AES 256 bit Encryption
Targeted Use Cases	• All Flash Arrays, Storage Servers, HPC, AI Training Servers • Needed for very low latency tiering, caching, write buffering, metadata storage, checkpointing • Needed for AI/ML algorithm processing

CXL 2.0 Support for Persistent Memory

- **PCIe enumeration**
  - NFIT not used for CXL
  - Leverage PCIe framework

- **MMIO registers**
  - Mailbox interface, etc.

- **Command Interface**

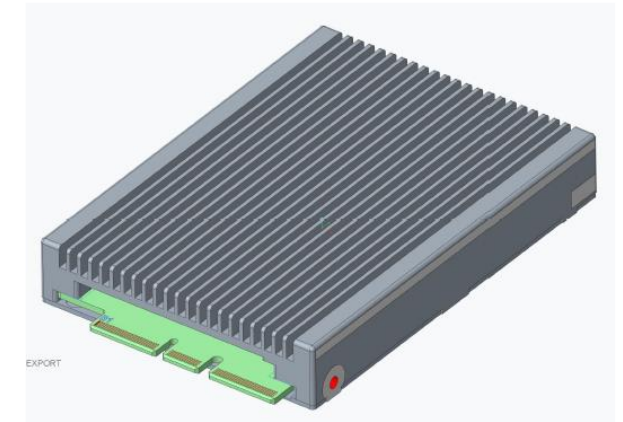- **Minor changes to external specs like ACPI/UEFI**
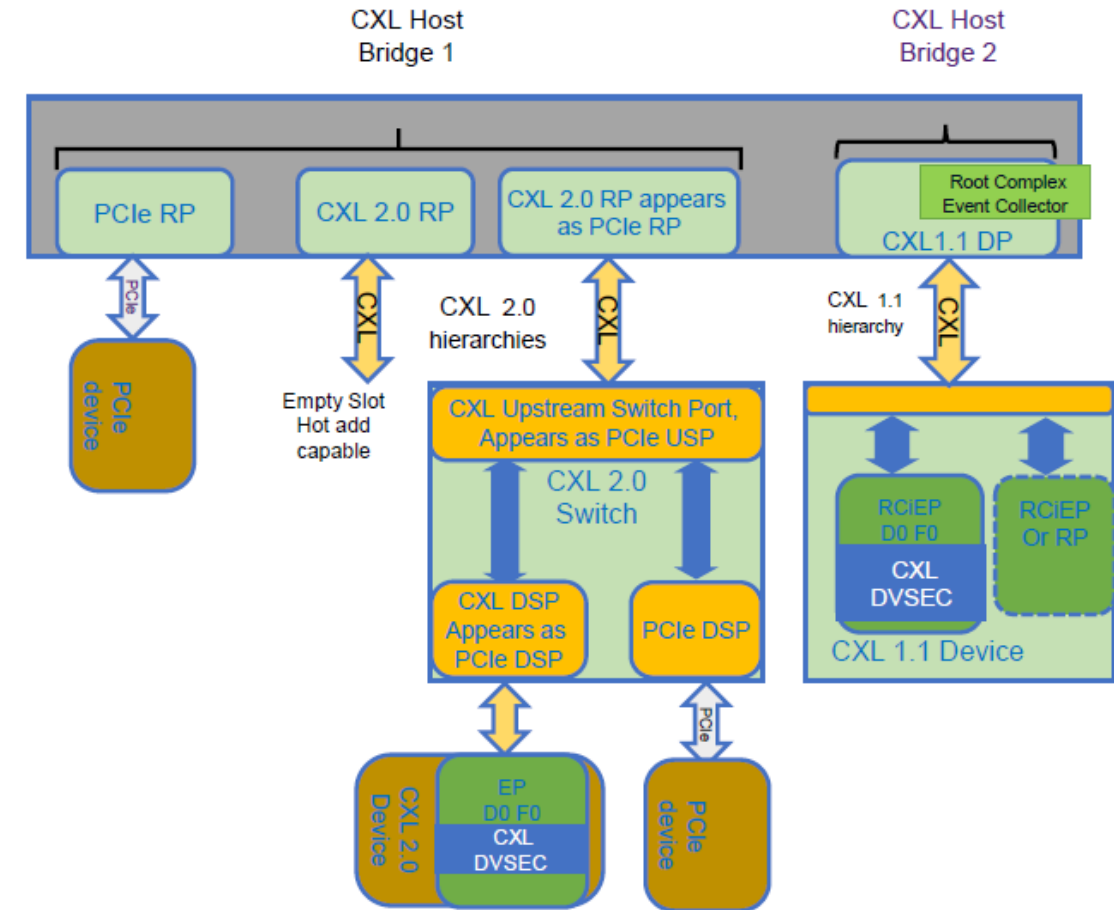
Source; SNIA PM Summit Andy Rudoff April 2021

CXL 2.0 Support for Persistent Memory

- Operate with standard CXL type 3 Linux and Windows memory driver for enumeration, configuration, I/O, and management.
- DRAM equivalent latency and endurance; Endurance limited only by Save/Restore cycles to NV
- Architecture supports multiple form factors
- End-to-End data protection from CXL host through Memory Buffer, NVC, and on all storage devices.
- NV data at rest must be encrypted.
- Support on module (device managed) or off module (host managed) NV Save energy source.

PERSISTENT MEMORY
+ SUMMIT 2022
COMPUTATIONAL STORAGE

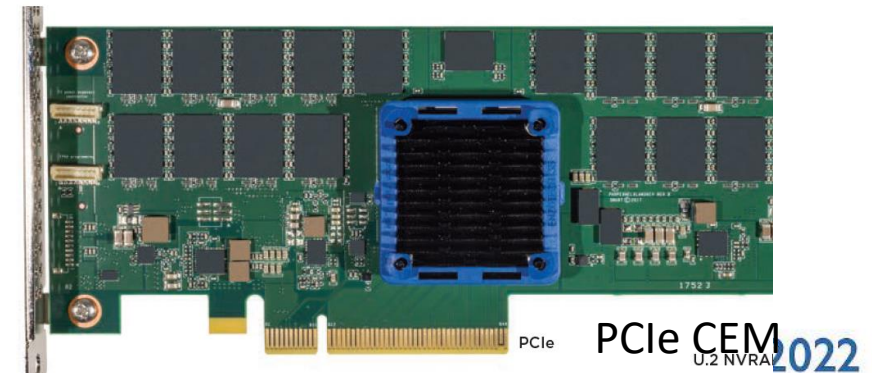# CXL™-based Persistent Memory - NV-XMM

- **CXL Type 3 PM (NV-XMM) Design Goals:**
  - On board power source for back-up power
  - Self managed back-up and restore
  - Standard Programming model for CXL Type-3 devices
  - DRAM equivalent latency and endurance;
  - Unlimited Endurance as active data writes directly go to DRAM
  - Encrypting data during back-up.

E3.S 2T

- **CXL Type 3 PM (NV-XMM)**
  - Memory Buffer to be used with RCD.
  - Infrequent Save/Restore cycles (< 10 per day)

E1.S
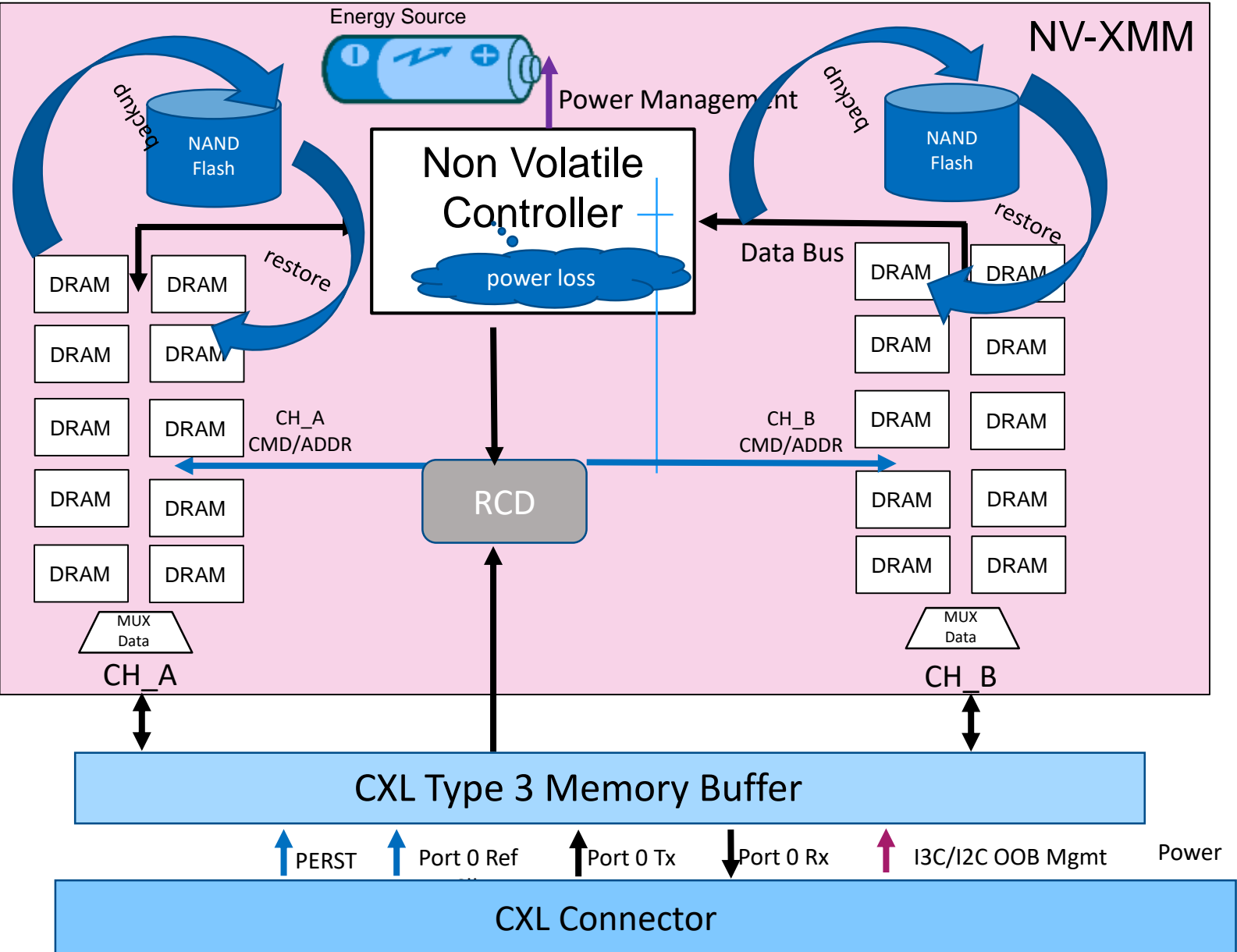
- **Form Factors**
  - E3.S / AIC / E1.S

PCIe CEM

CXL™ NV-XMM Architecture

Feature	Description
Memory	• DDR5
Protocol	• CXL 2.0
Form Factors	• E3.S 2T, AIC
Features	• Unlimited endurance • Latency <100ns • Secure backup
Targeted Use Cases	• All Flash Arrays, Storage Servers, • HPC, AI Training Servers • For very low latency tiering, caching, write buffering, metadata storage, checkpointing • For AI/ML algorithm processing • New applications…

CXL™ Adds Capacity, Bandwidth and Persistent Memory

Current Gen Systems
DDR4 NVDIMM-N
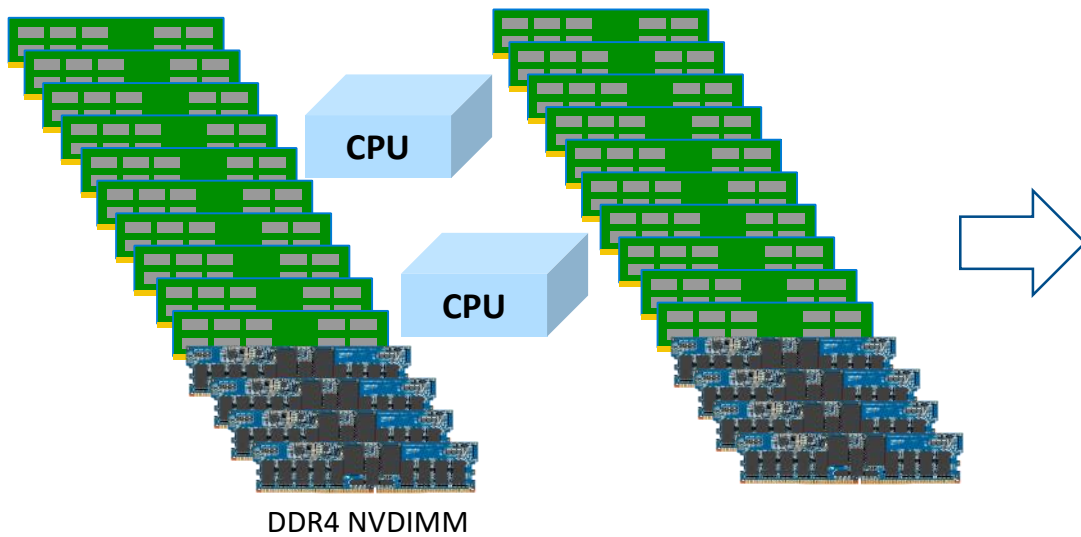
Next Gen Systems
DDR5 NVDIMM-N and CXL NV-XMM

CPU

CPU

DDR4 NVDIMM

CPU

CPU
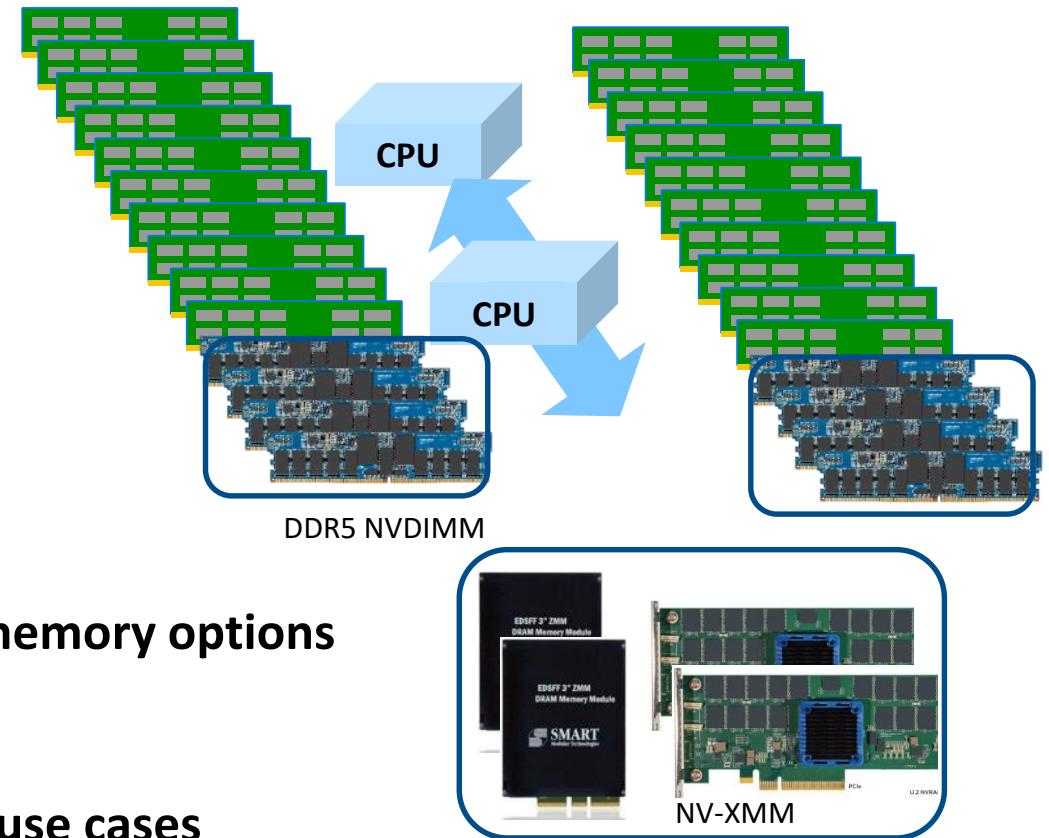
DDR5 NVDIMM

- **Expanded persistent memory options**
- **DDR5 NV-DIMM**
- **CXL NV-XMM**
- **New applications and use cases**

NV-XMM

Key Takeaways

- NVDIMM-N architecture transition from DDR4 to DDR5 without significant changes.

- Any latency-sensitive data that is continuously changing can benefit from NVDIMM and NV-XMM.

- DDR5 NVDMM-N is near-term and will fill a gap for high performance, low latency persistent memory applications.

- CXL NV-XMM will follow and also be used for low latency-sensitive applications.

- CXL NV-XMM will coexist with DDR5 NVDIMM-N

- CXL NV-XMM will open up new use cases for higher capacity and composability

PERSISTENT MEMORY
SNIA + SUMMIT 2022
COMPUTATIONAL STORAGE

Please take a moment to rate this session.

Your feedback is important to us.