STORAGE DEVELOPER CONFERENCE
SDC 21
BY Developers FOR Developers
Virtual Conference
September 28-29, 2021
A SNIA Event

Future of Storage Platform Architecture

Mohan J Kumar, Intel Fellow, Intel Corporation
Anjaneya "Reddy" Chagam, Cloud Architect, Intel Corporation

Notices & Disclaimers

Intel technologies may require enabled hardware, software or service activation. Your costs and results may vary.

No product or component can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. For more complete information about performance and benchmark results, visit http://www.intel.com/benchmarks .

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/benchmarks .

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

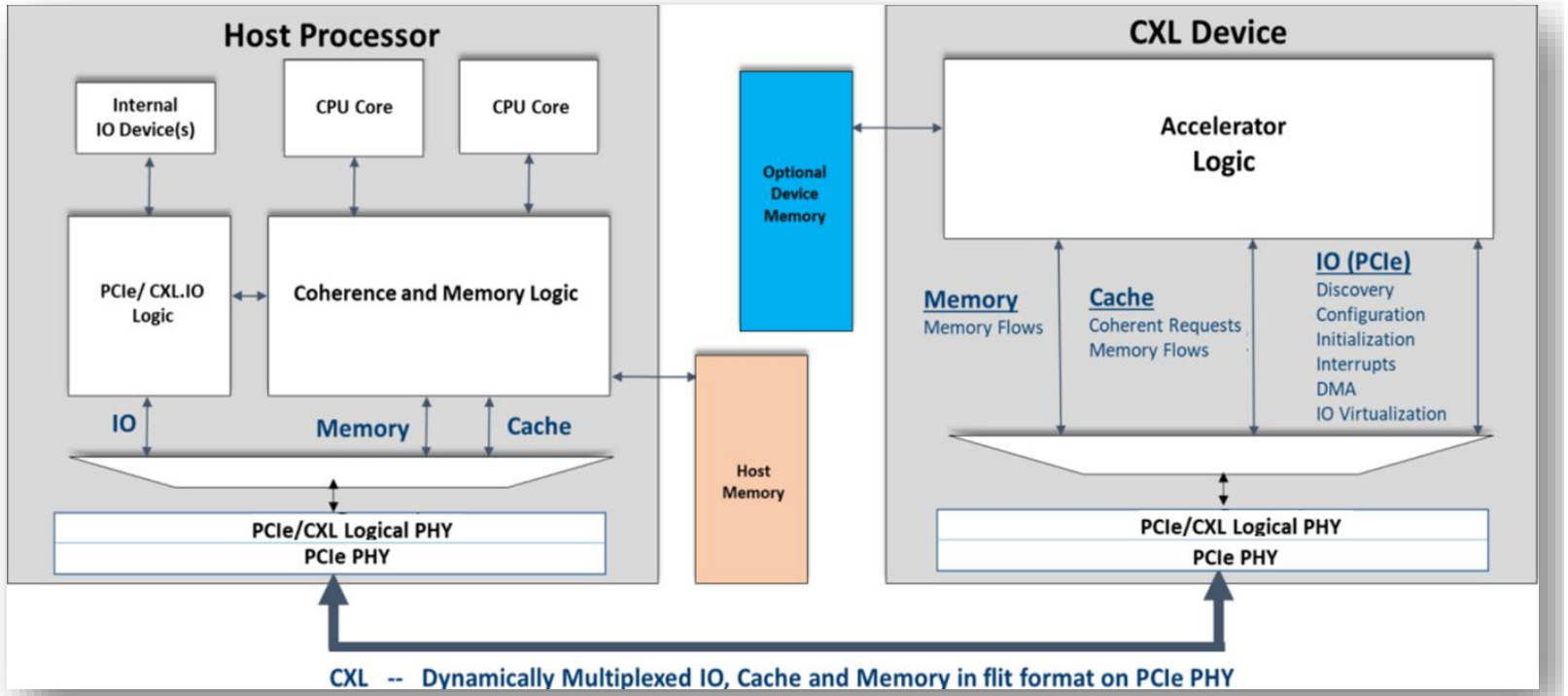
Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

.

Agenda

- CXL Overview
- CXL for Memory/Storage converged architecture
- Future Storage Architecture
- Summary & Next Steps

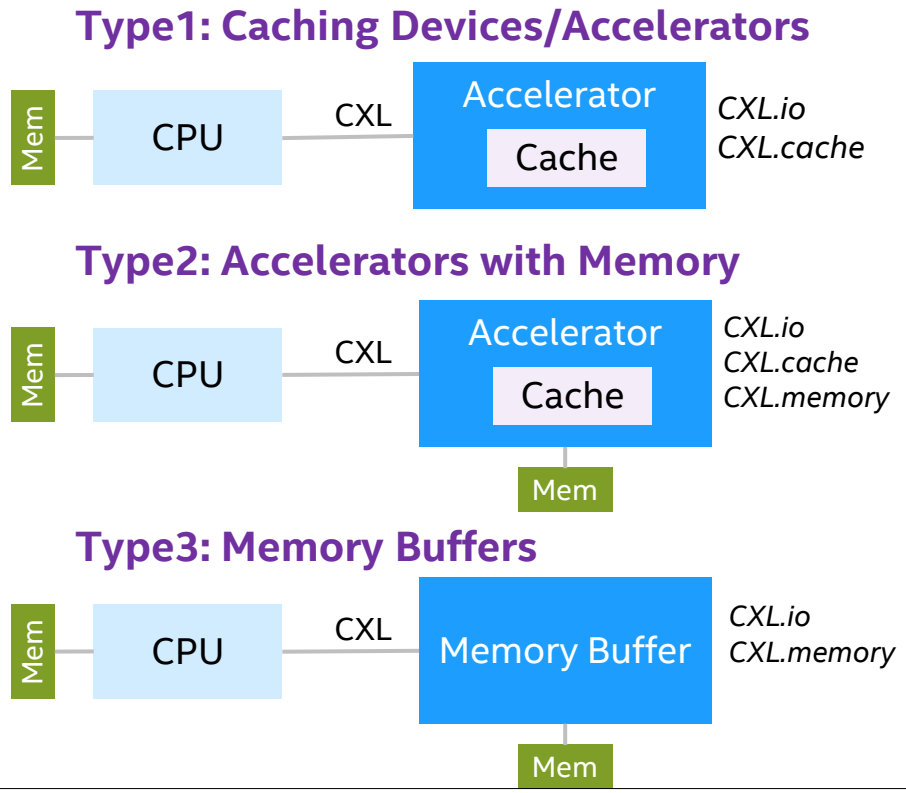
CXL Overview



Source: Compute Express Link™ (CXL™) specification

- **Open industry standard processor interconnect**
- **Alternate protocol that runs on standard PCIe PHY**
- **3 sub-protocols multiplexed on single link**
- **High-bandwidth, low-latency**
- **Coherent interface**

CXL for Memory/Storage Converged Architecture

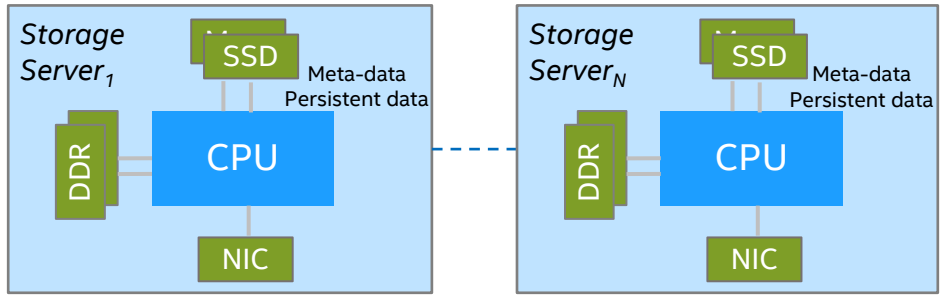


- **Memory Bandwidth/Capacity expansion**
- **Resource Pooling - Memory/Accelerator**
- **Disaggregation**
- **Computational Offloads**
  - AI, HPC, Comms

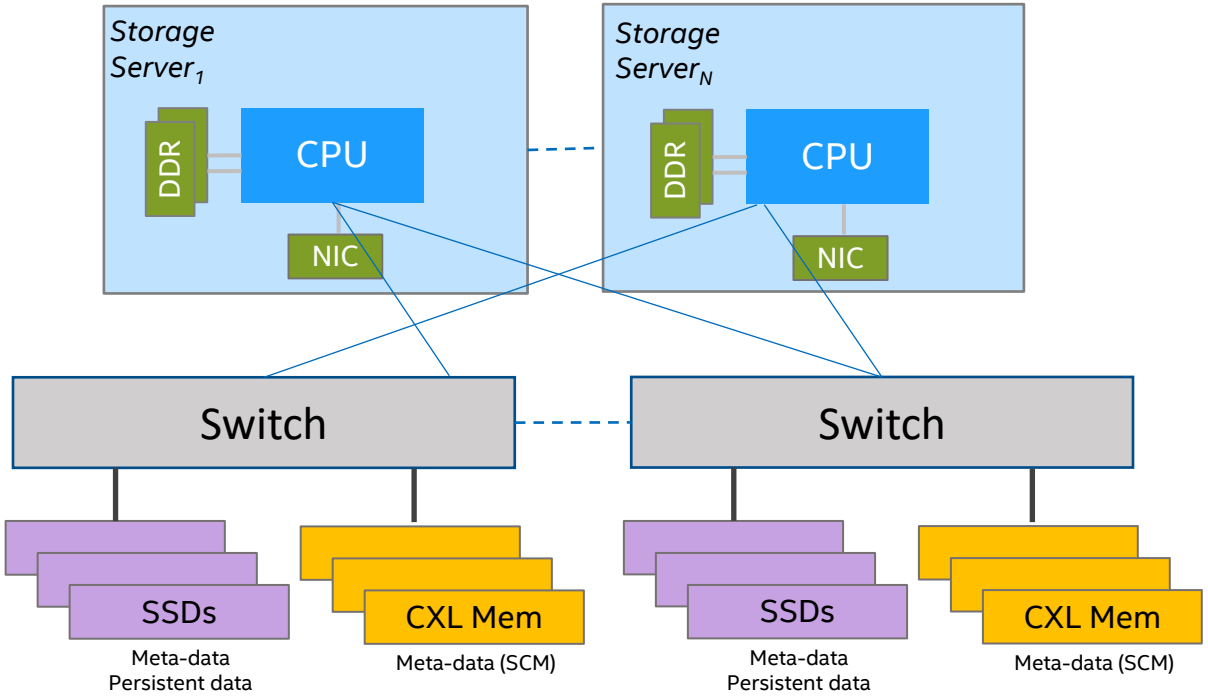
Storage Platform Architecture Opportunities

- High Available architecture for scale out storage using CXL
- Easier Memory persistence to speed up SDS meta-data operations
- Storage/Memory converged device unlocks potential for memory centric usages
- CXL accelerators provide computational offload architecture for storage

Storage Architecture – SDS High Availability

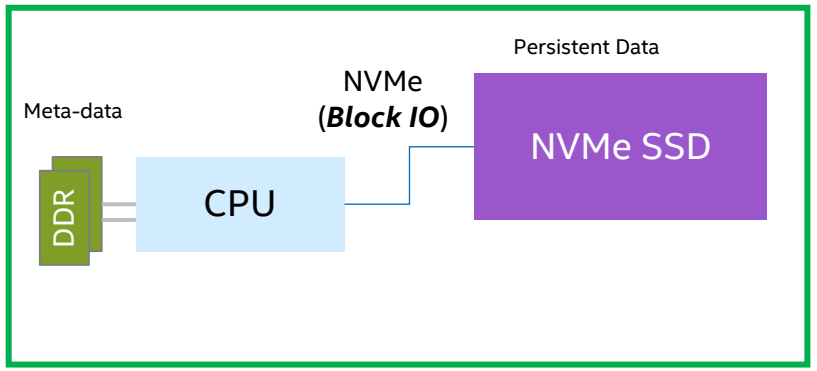


- **SDS - shared nothing architecture**
- **Any node failure triggers cluster wide rebuild/rebalancing**
- **Takes several hours rehydrate data from peer storage cluster nodes and rebuild meta-data**

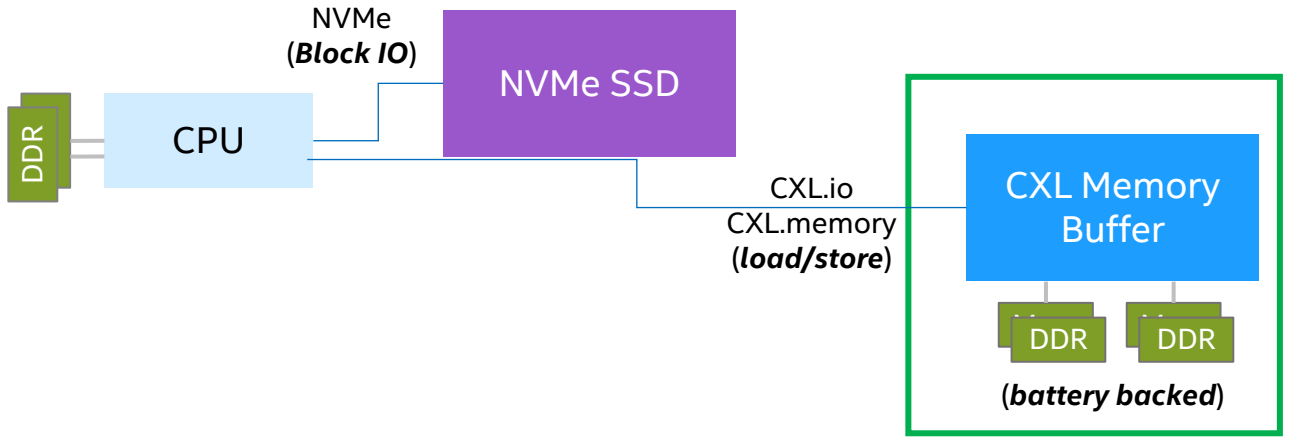


- **Host failure doesn't trigger cluster wide rebuild**
- **Meta-data in SCM helps in reducing rebuild time for data**

Storage Architecture – Persistence



Persistency Domain

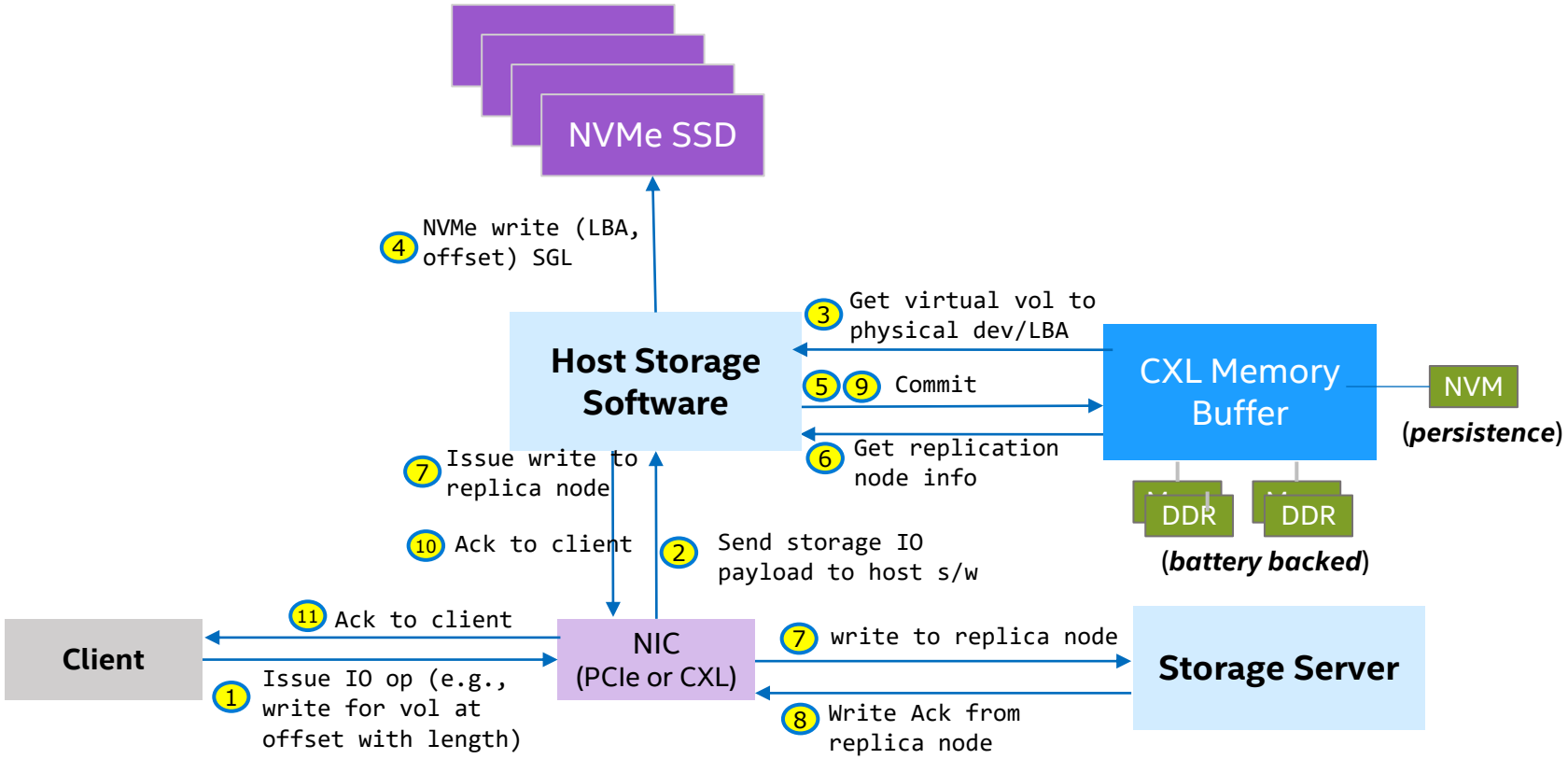


Persistency Domain

- Full System persistence required to keep Meta data in memory
- Persistence is Platform/CPU dependent
- Software changes required for SCM persistence
- Block IO persistence is slow for meta-data

- Easier to implement persistence for CXL Memory (e.g., battery backed for power fail)
- Removes Platform / CPU dependence
- Speeds up meta-data operations with DDR latency while achieving persistence

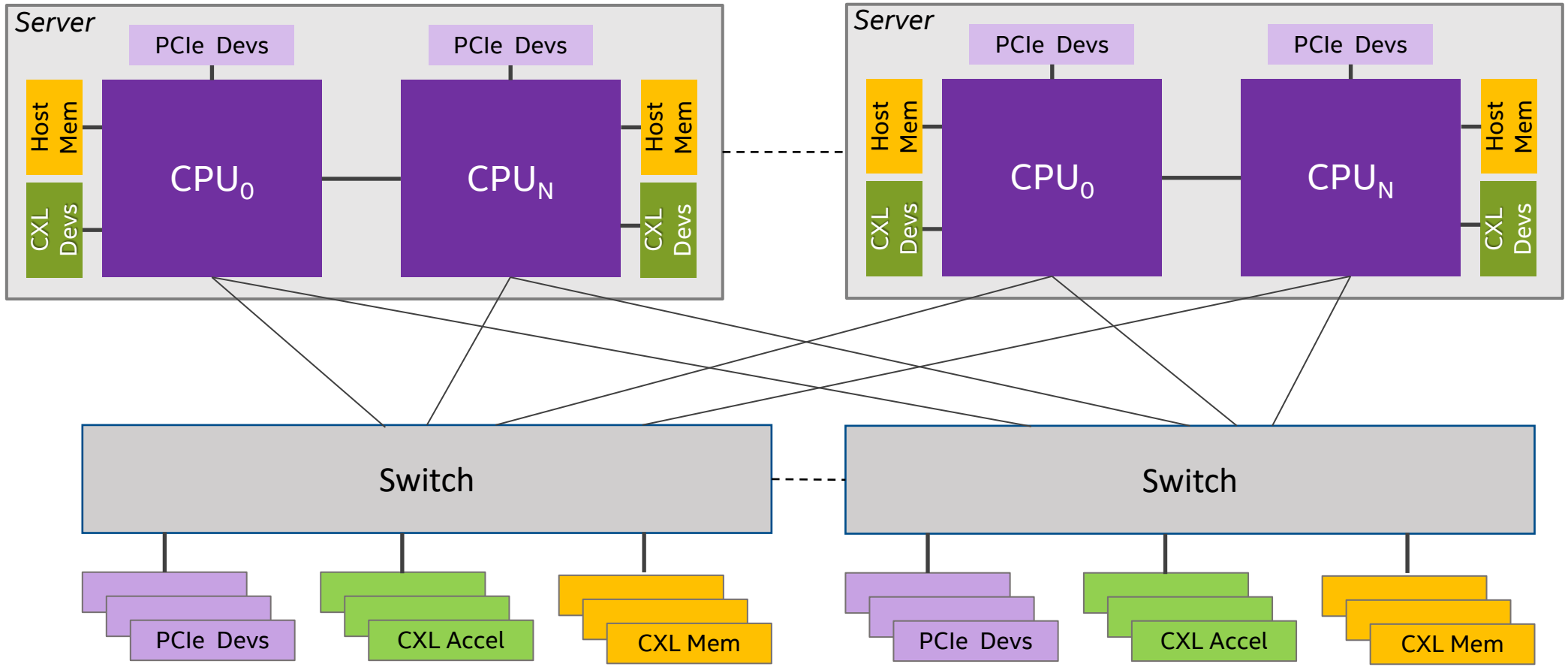
Storage Architecture – SDS Metadata Acceleration



- Mounts network block device
- Issues block io (read/write etc.)

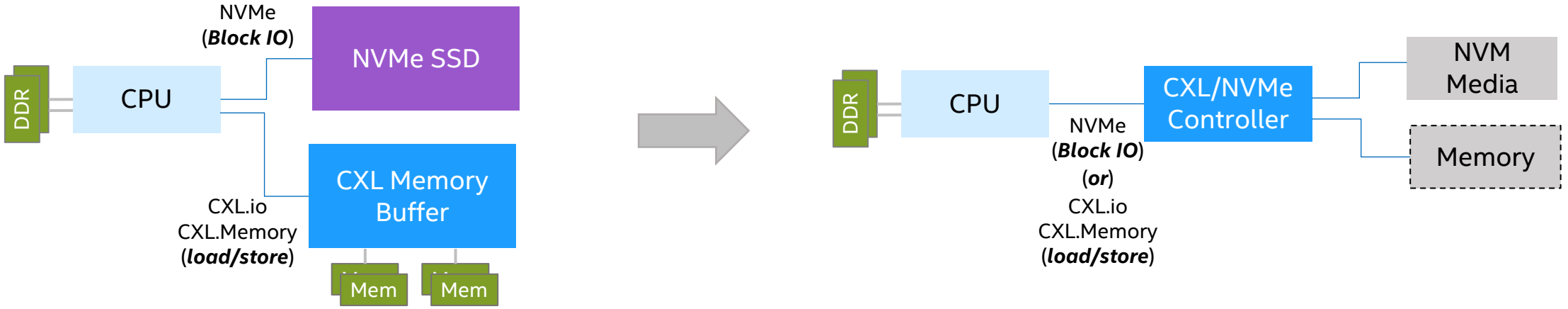
- **Avoids complex data protection implementation in software (e.g., transaction logs, rollbacks)**
- **Meta-data is write intensive – reduces latency significantly**

CXL for Memory/Storage converged architecture



- **Disaggregation and Resource Sharing**
- **Memory hierarchy (Persistent/Volatile)**
- **P2P Communication**

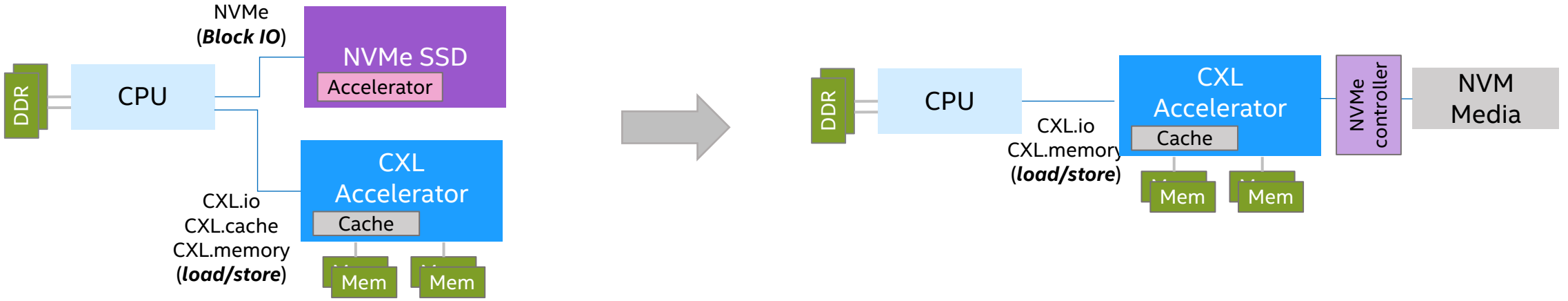
Storage Architecture – Converged Device



- **Monolithic - Protocol & Media**
- **Device training/RAS – Unique to device type**
- **Software – block v/s memory semantics**

- **Protocol is late binding decision based on workload requirements**
- **Converged device training/Platform RAS flows**
- **Open new usages for SSDs – Memory semantics**

Storage Architecture – Computational Offloads



- **Compute offload semantics – protocol dependent**
- **Computational storage – early stages of maturity**
- **Software changes – intrusive**

- **CXL for computational storage offloads**
- **Enable memory and SCM semantics for persistence**
- **Dual protocol support for backward compatibility**

Summary and Next Steps

- CXL opens opportunities for storage architecture innovations
- Storage/Memory converged device unlocks potential for memory centric usages
- CXL accelerators provide computational offload architecture for storage
- CXL device DDR persistence to speed up SDS meta-data operations
- High Available architecture for scale out storage using CXL constructs to reduce cluster wide rebuild/recovery time

Please take a moment to rate this session.

Your feedback is important to us.

