

STORAGE DEVELOPER CONFERENCE



Fremont, CA
September 12-15, 2022

BY Developers FOR Developers

A SNIA[®] Event

Why Should the Storage Community Care about CXL?

YANG SEOK KI, Ph.D.

Vice President, Memory Solutions Lab., Samsung Electronics

SAMSUNG

SAMSUNG

Outline

- Why Is CXL So Hot?
- What Is CXL Really?
- Convergence in Action
 - Memory perspective
 - Storage perspective
 - Computing perspective
- Memory Technology Continuum
- Conclusion

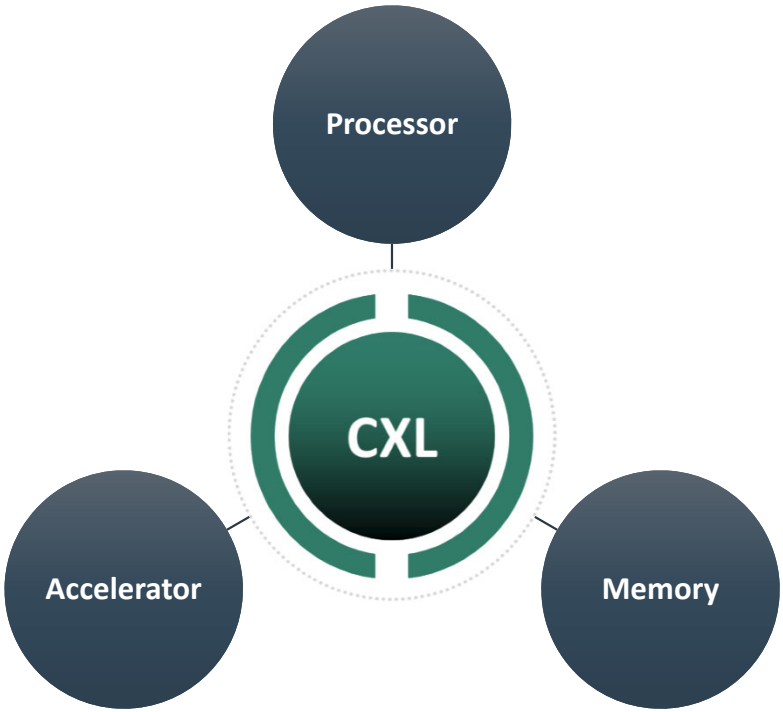




Why Is CXL So Hot?

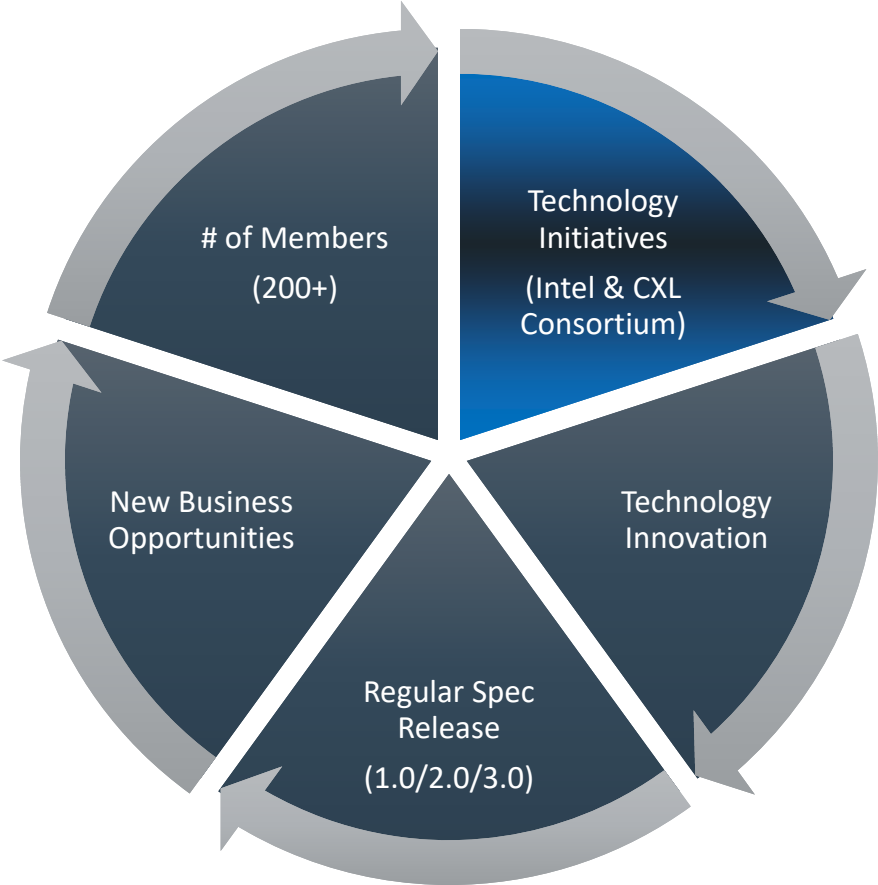
CXL (Compute Express Link)

- **CXL is an industry-supported cache-coherent interconnect for processors, memory expansion and accelerators** [source: www.computeexpresslink.org]



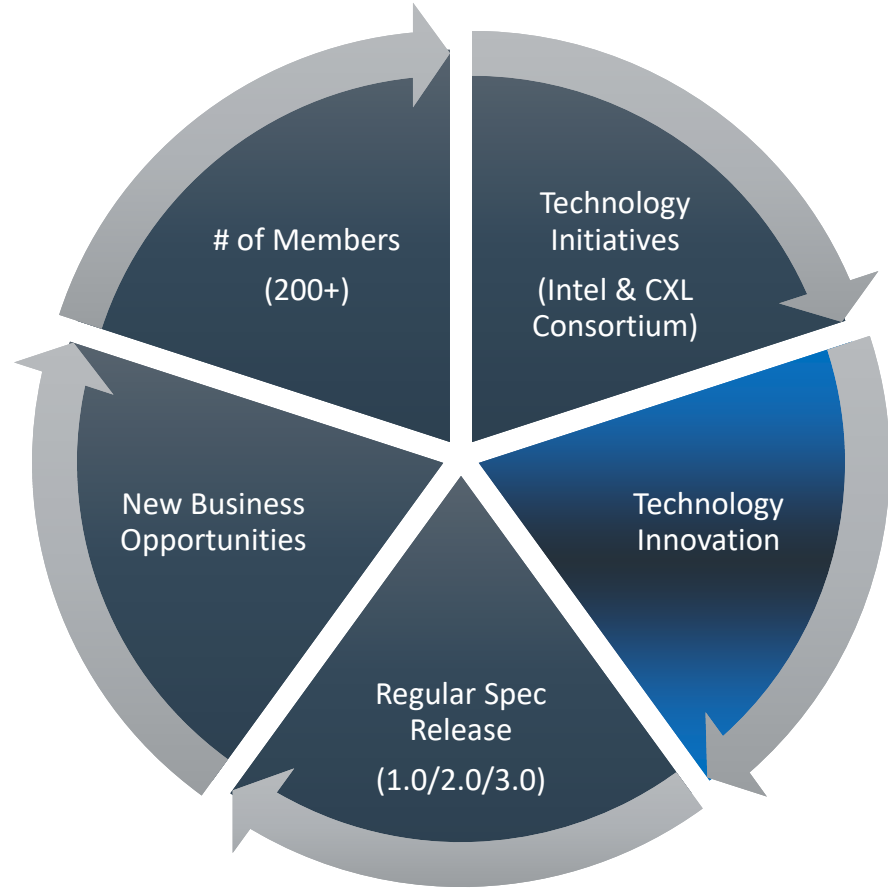
Technological Momentum

- A virtuous cycle is created



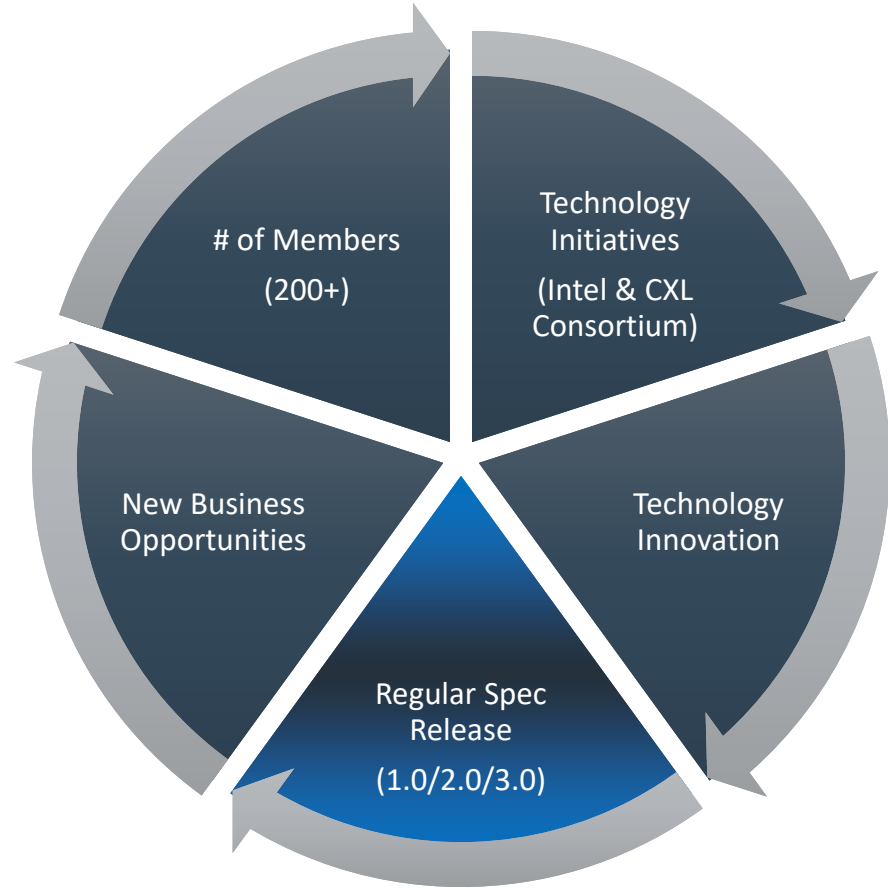
Technological Momentum

- A virtuous cycle is created



Technological Momentum

- A virtuous cycle is created



Continual Evolution

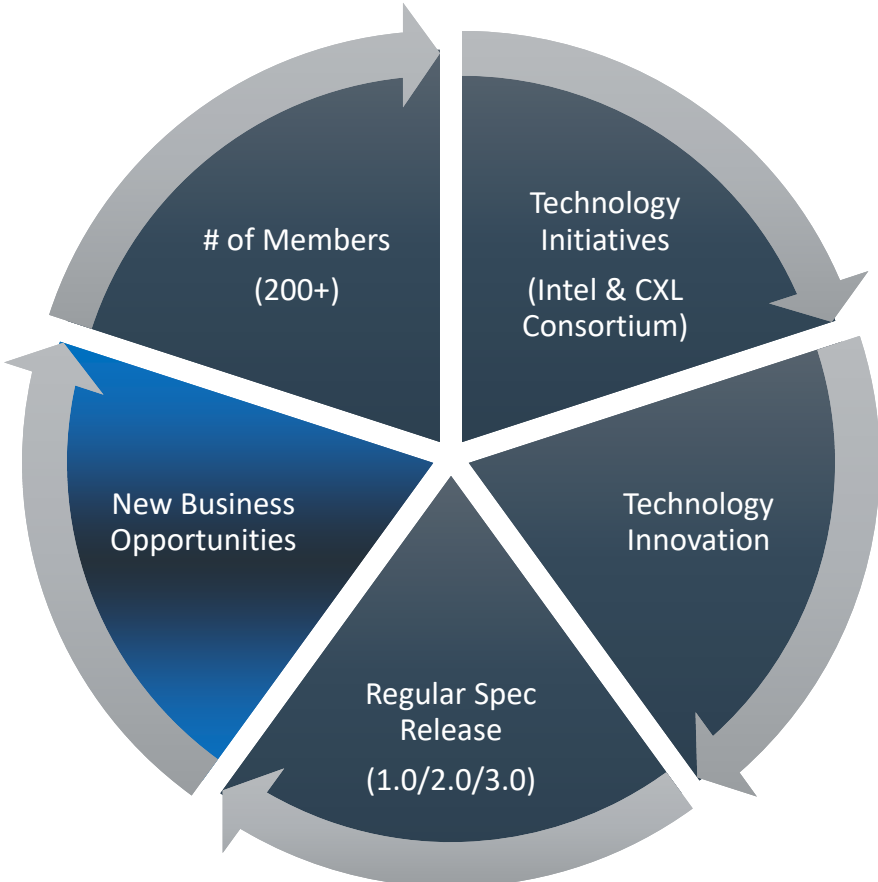
- Regular spec updates strongly supported by 200+ companies

Features	CXL 1.0 / 1.1	CXL 2.0	CXL 3.0
Release date	2019	2020	1H 2022
Max link rate	32GTs (PCIe 5.0)	32GTs (PCIe 5.0)	64GTs (PCIe 6.0)
Flit 68 byte (up to 32 GTs)	✓	✓	✓
Flit 256 byte (up to 64 GTs)			✓
Type 1, Type 2 and Type 3 Devices	✓	✓	✓
Memory Pooling w/ MLDs		✓	✓
Global Persistent Flush		✓	✓
CXL IDE		✓	✓
Switching (Single-level)		✓	✓
Switching (Multi-level)			✓
Multiple Type 1/Type 2 devices per root port			✓
Memory sharing (256 byte flit)			✓
Symmetric coherency (256 byte flit)			✓
Direct memory access for peer-to-peer			✓
Fabric capabilities (256 byte flit)			✓

Source: FMS'22

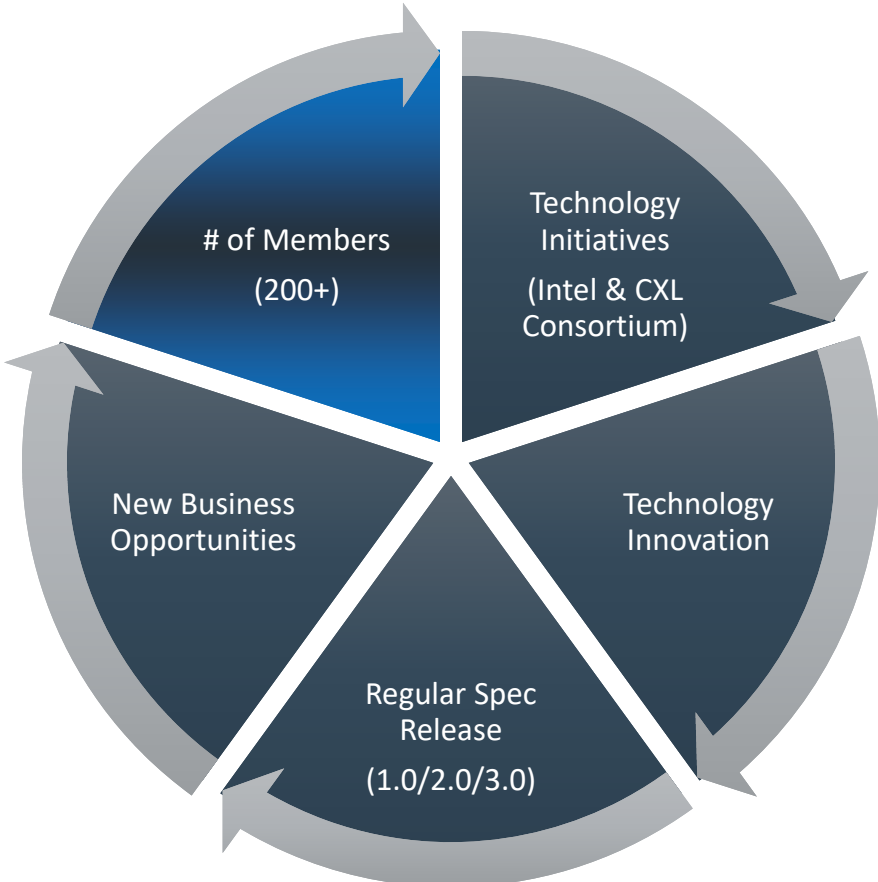
Technological Momentum

- A virtuous cycle is created



Technological Momentum

- A virtuous cycle is created

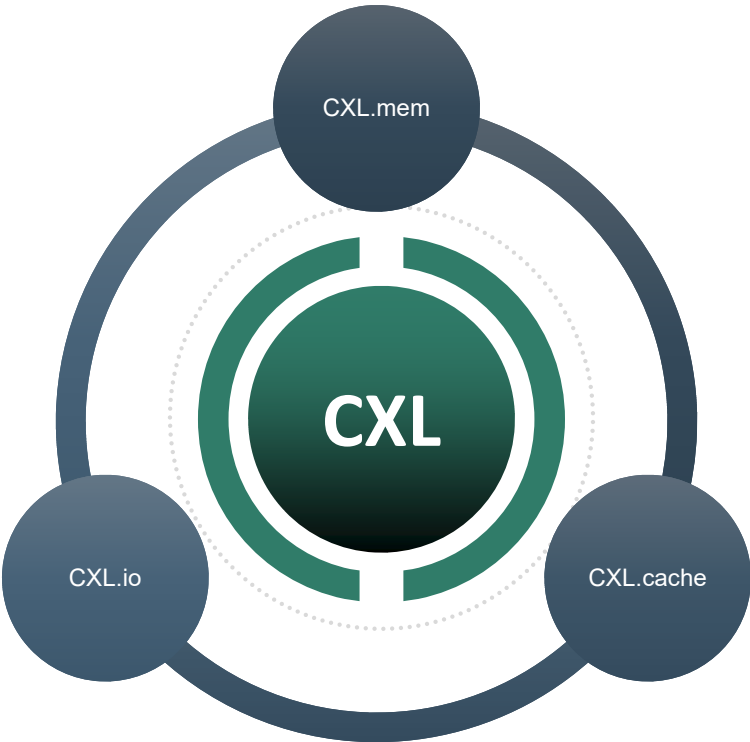




What Is CXL Really?

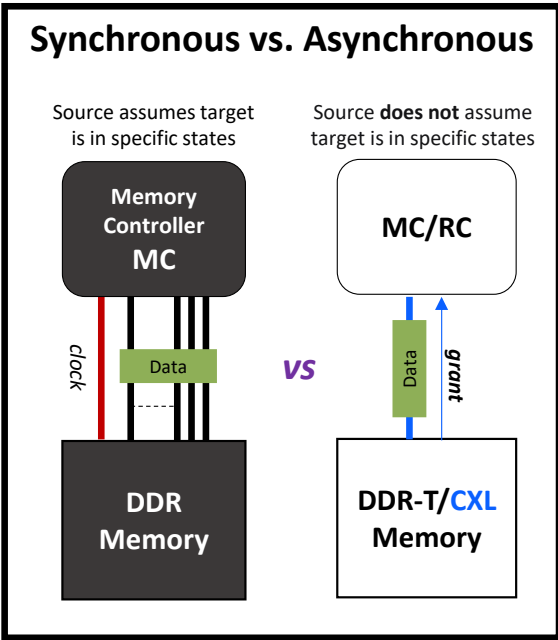
Additional Technical Definition of CXL

- CXL is an asynchronous blocking serial memory interface over variable latency fabrics, optionally supporting (a)symmetric coherency

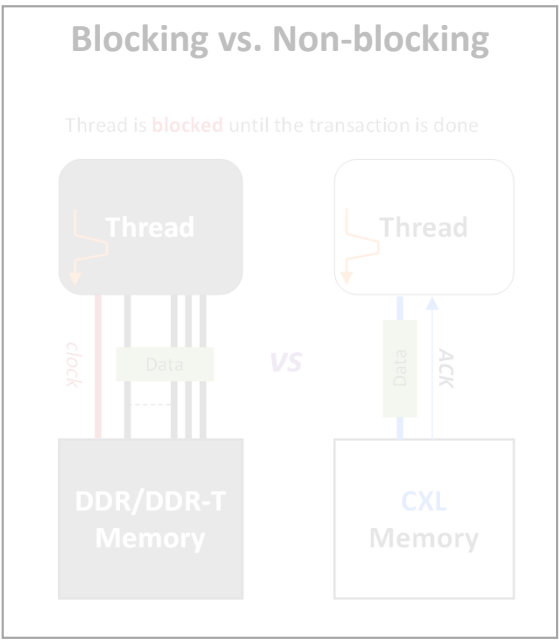


CXL from the Memory Interface Perspective

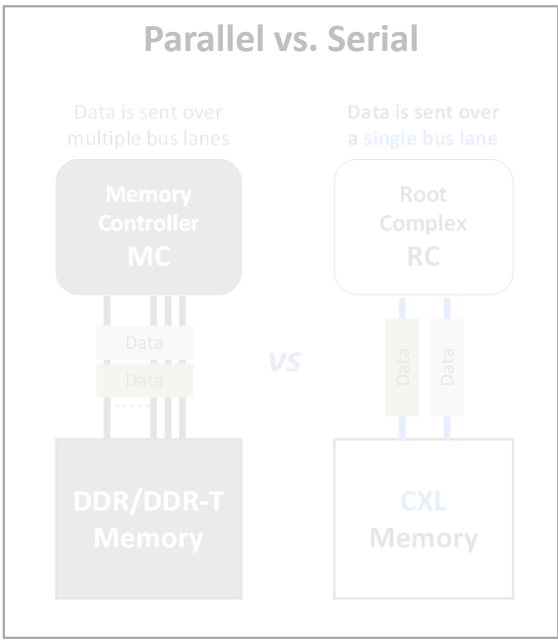
- CXL is an **asynchronous** blocking serial memory interface over variable latency fabrics, optionally supporting (a)symmetric coherency



→ Loose coupling of CPU and memory



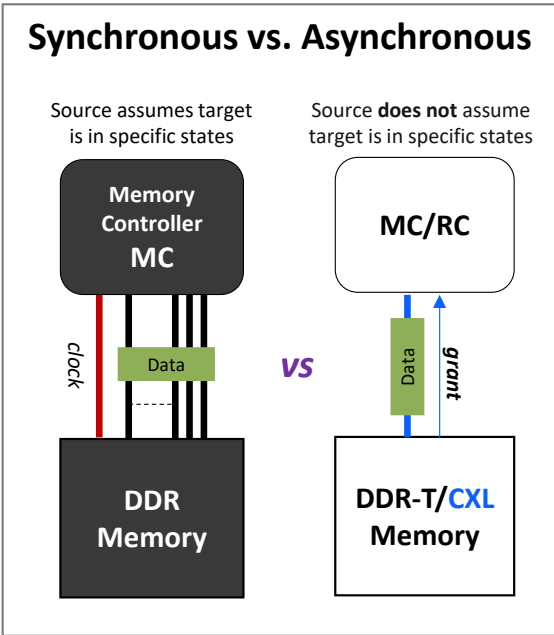
→ Low transaction complexity



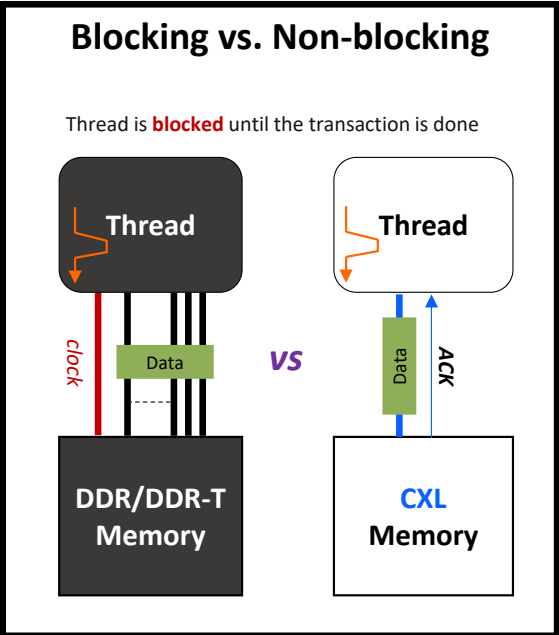
→ Longer distance high clock

CXL from the Memory Interface Perspective

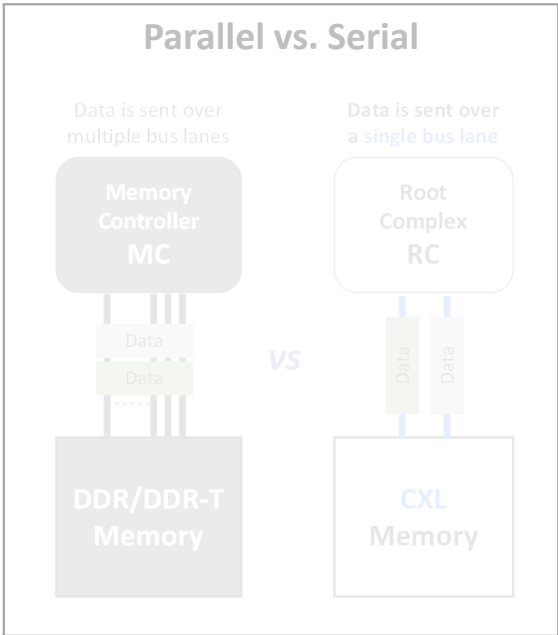
- CXL is an asynchronous **blocking** serial memory interface over variable latency fabrics, optionally supporting (a)symmetric coherency



→ Loose coupling of CPU and memory



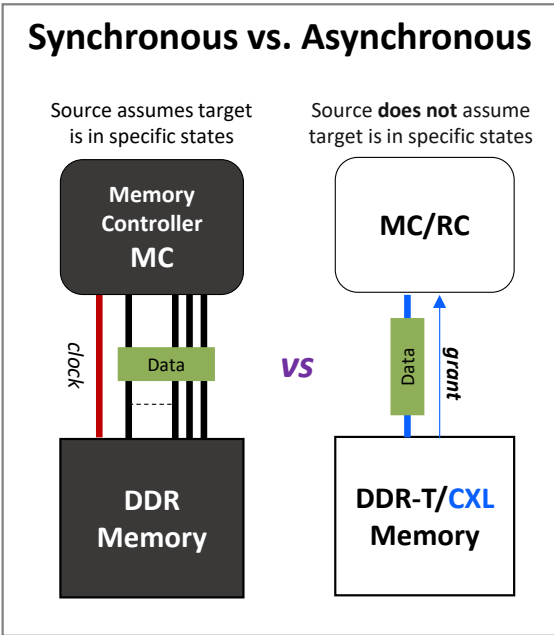
→ Low transaction complexity



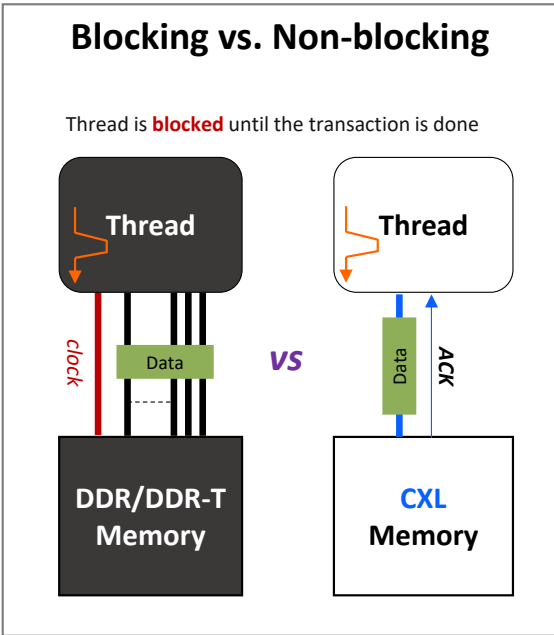
→ Longer distance high clock

CXL from the Memory Interface Perspective

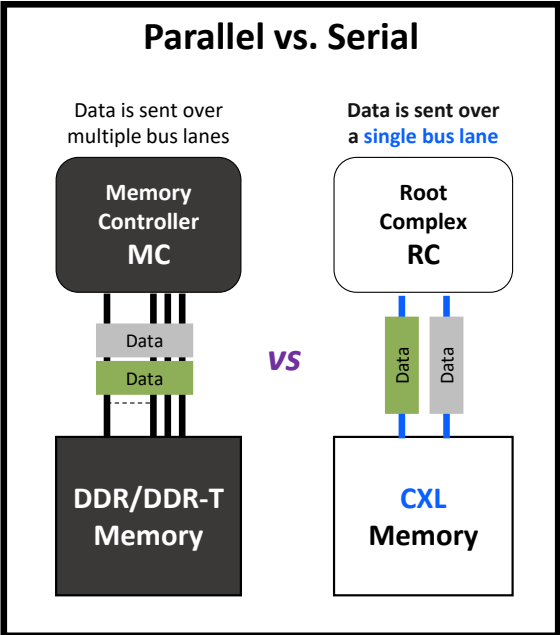
- CXL is an asynchronous blocking **serial** memory interface over variable latency fabrics, optionally supporting (a)symmetric coherency



→ Loose coupling of CPU and memory



→ Low transaction complexity

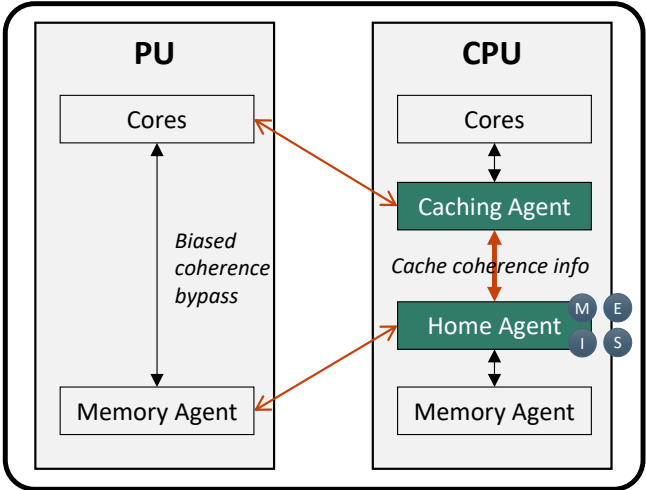


→ Longer distance high clock

CXL from the Cache Coherence Perspective

- CXL is an asynchronous blocking serial memory interface over variable latency fabrics, optionally supporting (a)symmetric coherency

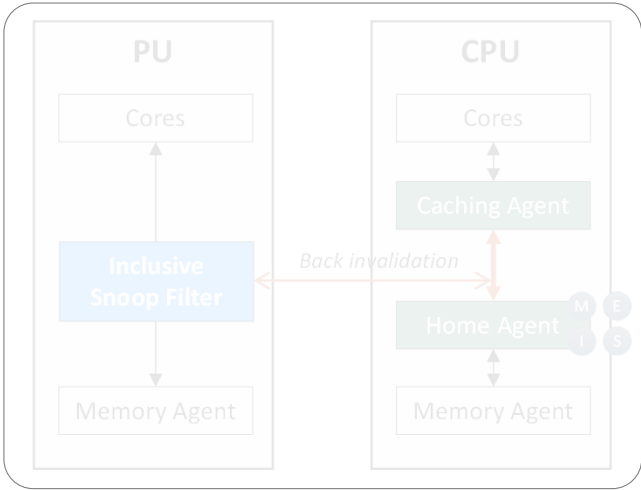
Asymmetric Cache Coherency Protocol
(CXL 1.x)



Source: Hot interconnect'19

→ Simple coherency resolution

Symmetric Cache Coherency Protocol
(CXL 3.0)



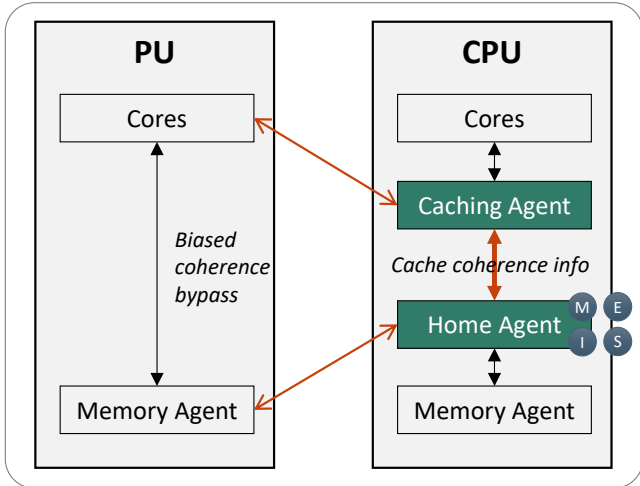
VS

→ P2P communication

CXL from the Cache Coherence Perspective

- CXL is an asynchronous blocking serial memory interface over variable latency fabrics, optionally supporting (a)symmetric coherency

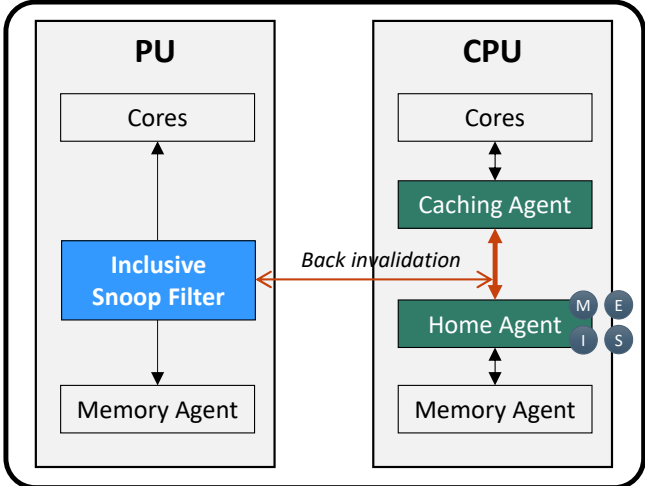
Asymmetric Cache Coherency Protocol
(CXL 1.x)



Source: Hot interconnect'19

→ Simple coherency resolution

Symmetric Cache Coherency Protocol
(CXL 3.0)

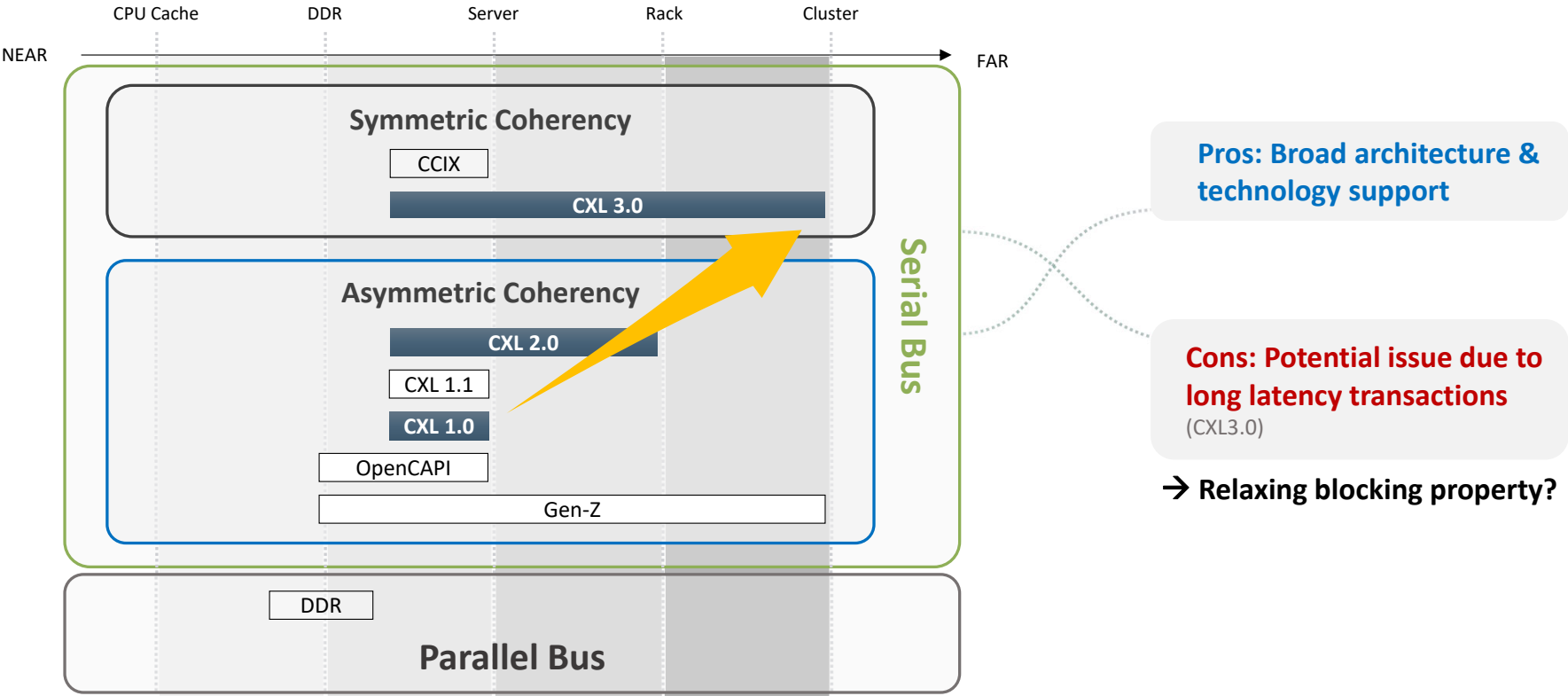


VS

→ P2P communication

CXL from the Access Latency Perspective

- CXL is an asynchronous blocking serial memory interface over **variable latency** fabrics optionally supporting (a)symmetric coherency

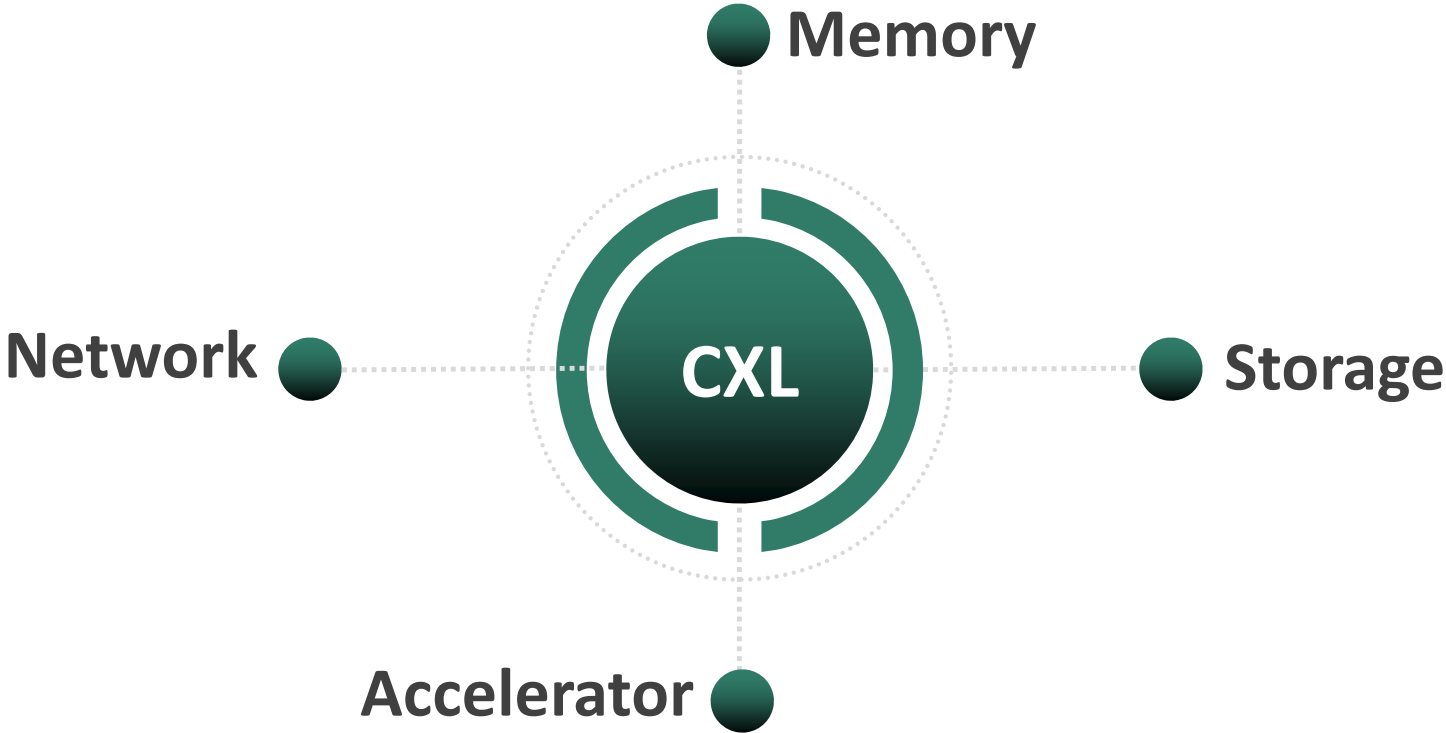




Convergence in Action

Convergence in Action

- Memory, storage, network and accelerator are converged through consistent interfaces

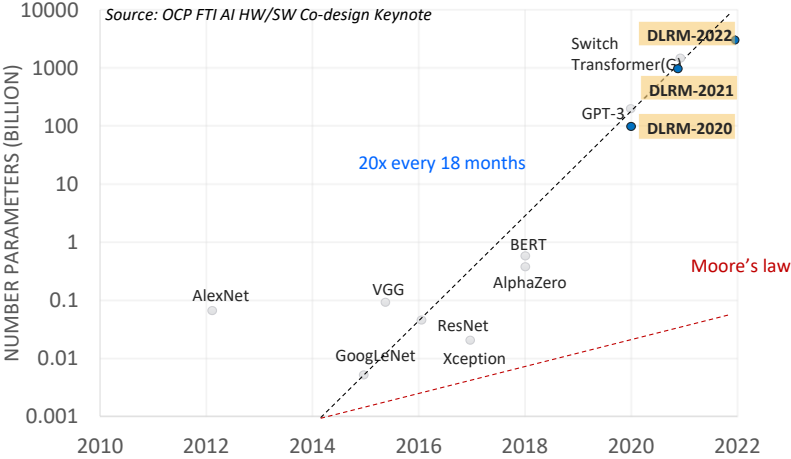
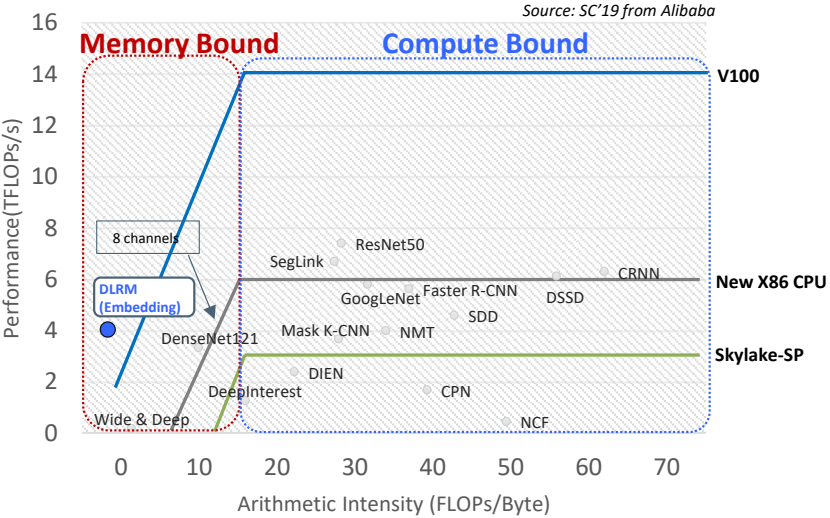




Convergence in Action: Memory Perspective

Memory Capacity and Bandwidth Challenge

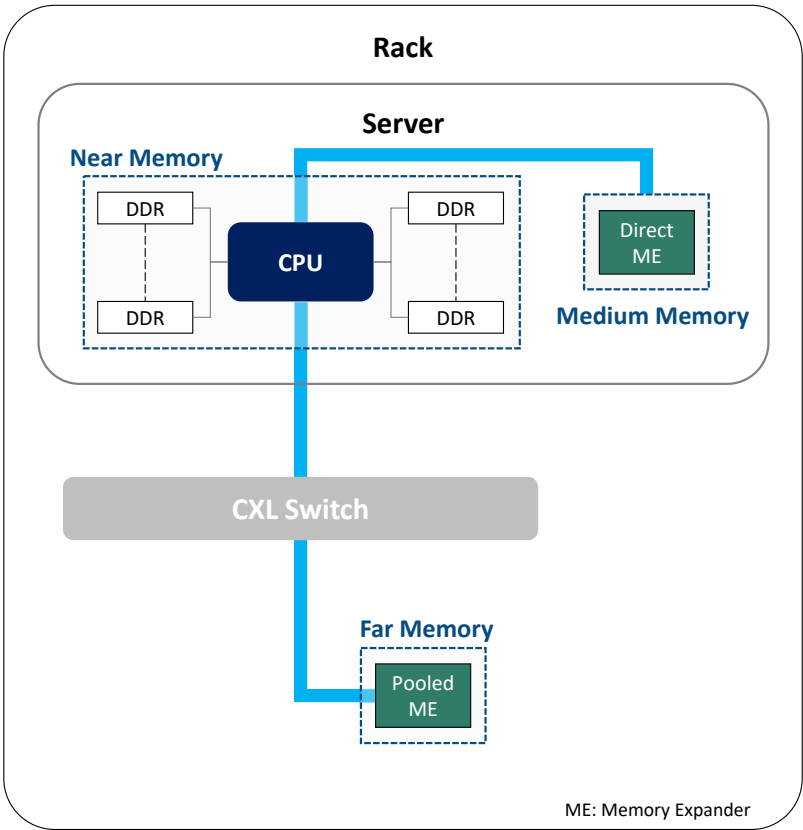
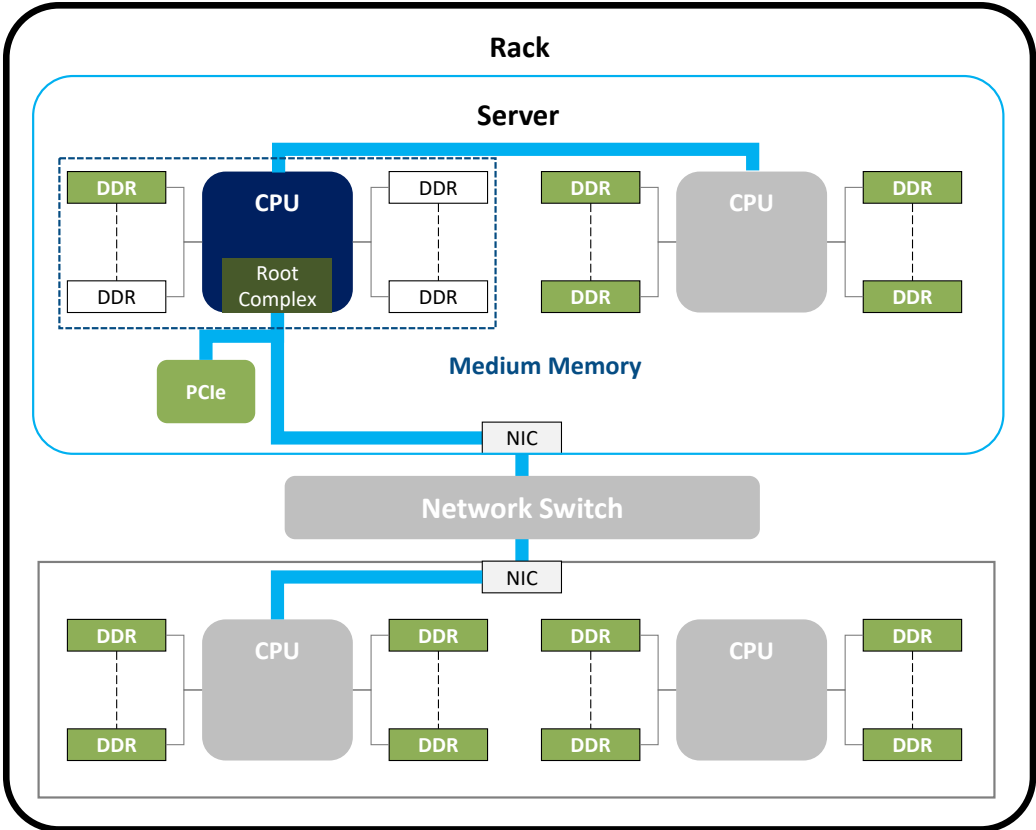
- Demanding more capacity and bandwidth than system can handle



CPU architecture optimized for typical workloads

Convergence of Memory Expansion Fabric into CXL

- Extended memory over DDR, processor interconnect, IO bus, and network into CXL-based memory



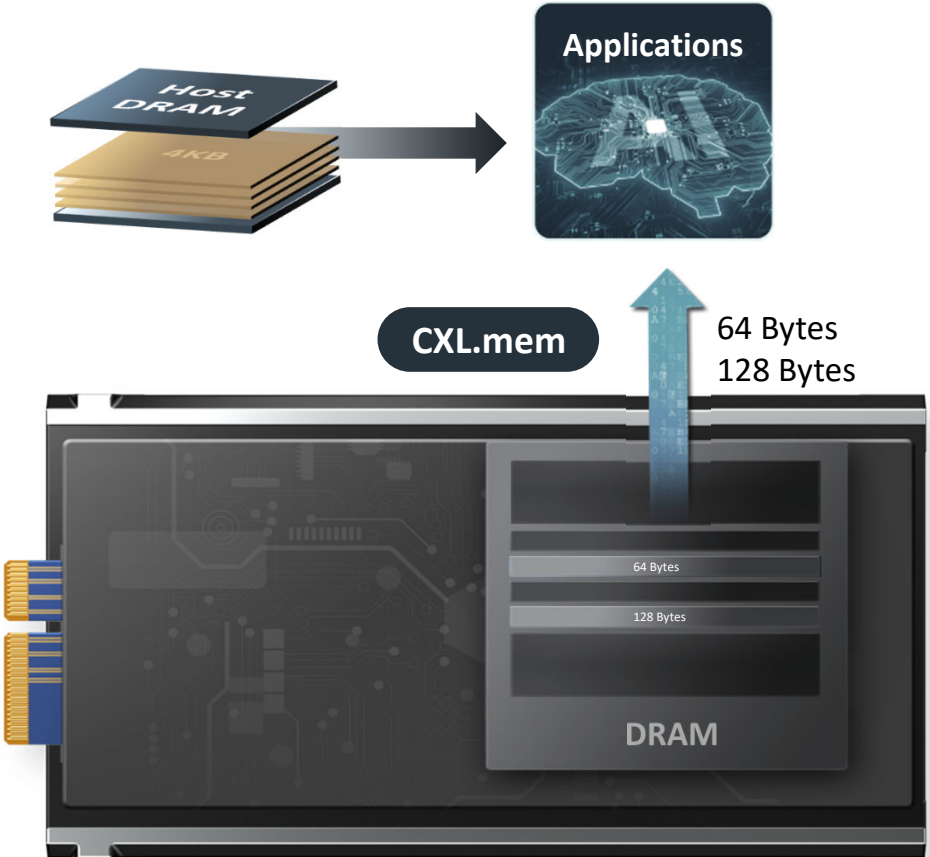
ME: Memory Expander

CXL-based Memory Expander Concept

- Loosely coupled memory expansion, operating outside of DDR

More Capacity

More Bandwidth



Samsung Memory Expander

- Industry 1st CXL-based memory expander product

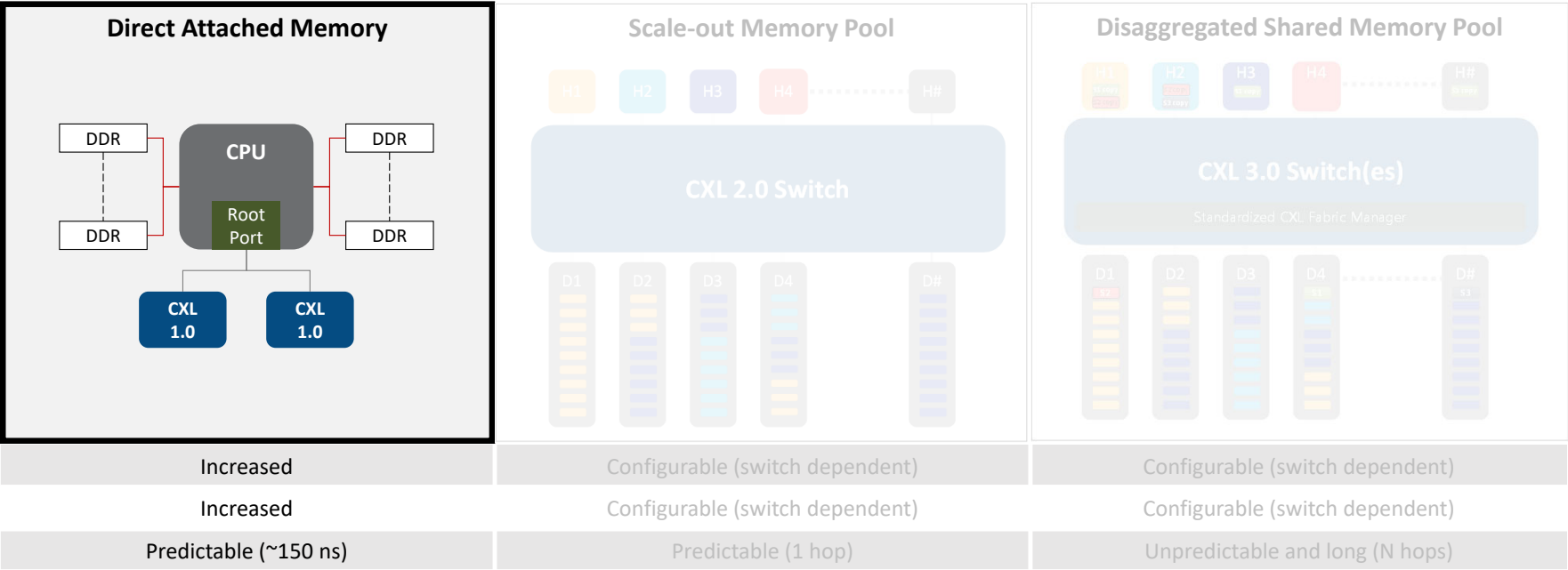


CXL-based Memory Expander	
Form Factor	E3.S 2T
Media	DDR5 4800
Capacity	Max 512GB
CXL Link	PCIe 5 X8
Max Bandwidth	32 GB/s
Availability	Q3/22
Version	CXL 2.0
Device Type	Type-3
RAS	<ul style="list-style-type: none">• Viral• Poisoning• Memory error injection• Multi-symbol ECC• Media scrubbing• Post package repairs (hard/soft)

Roughly close to the bandwidth of a single DDR channel with an average of 2-3x DDR latency

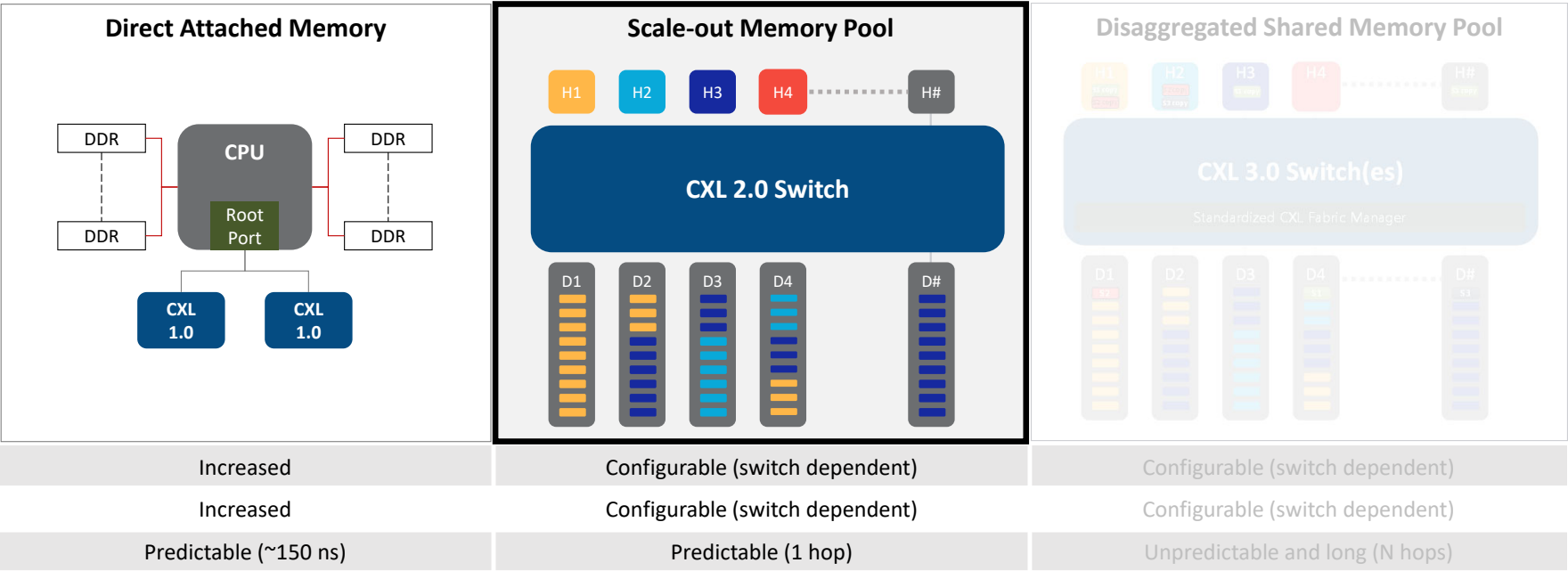
Memory Subsystem Architectures

- Toward disaggregated composable memory



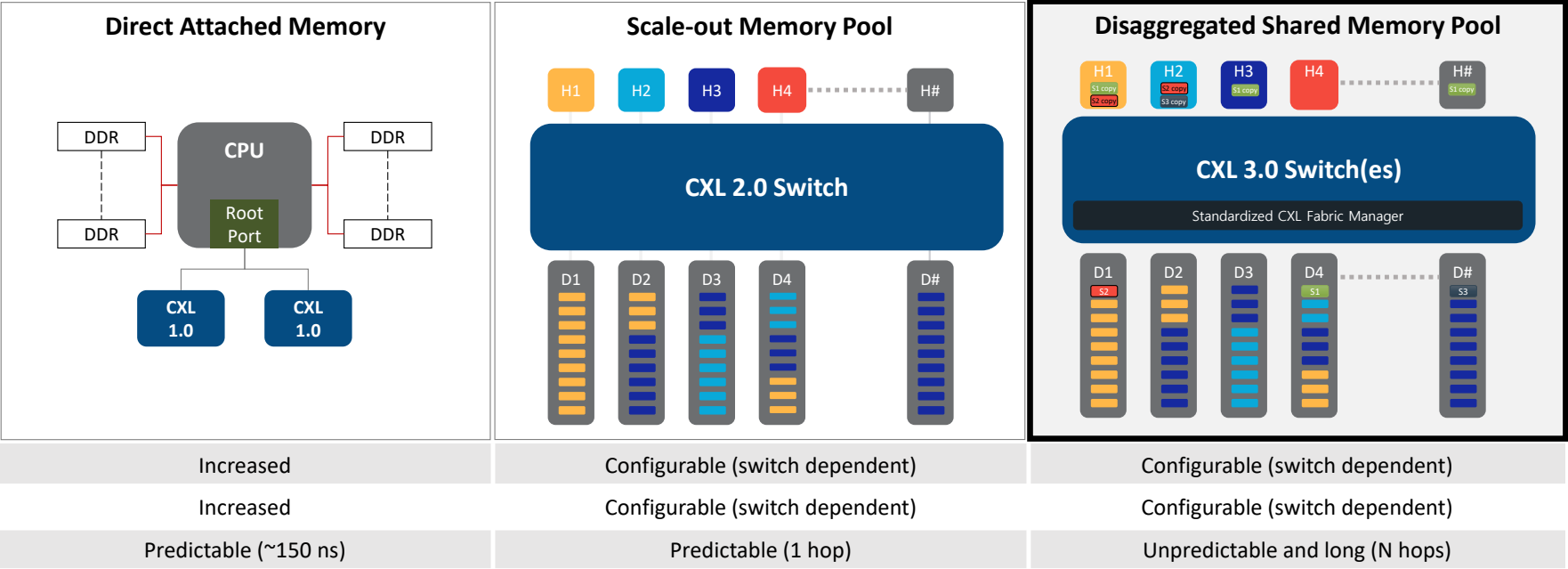
Memory Subsystem Architectures

- Toward disaggregated composable memory



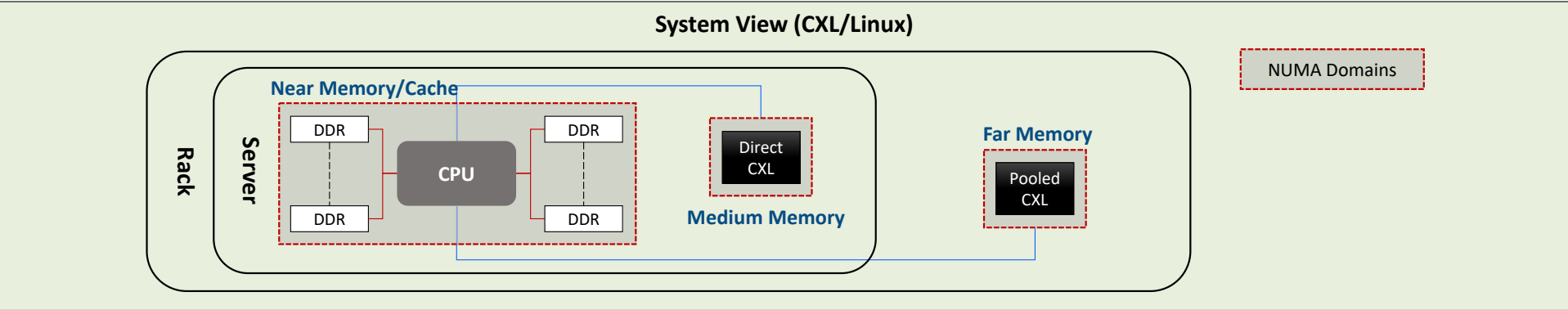
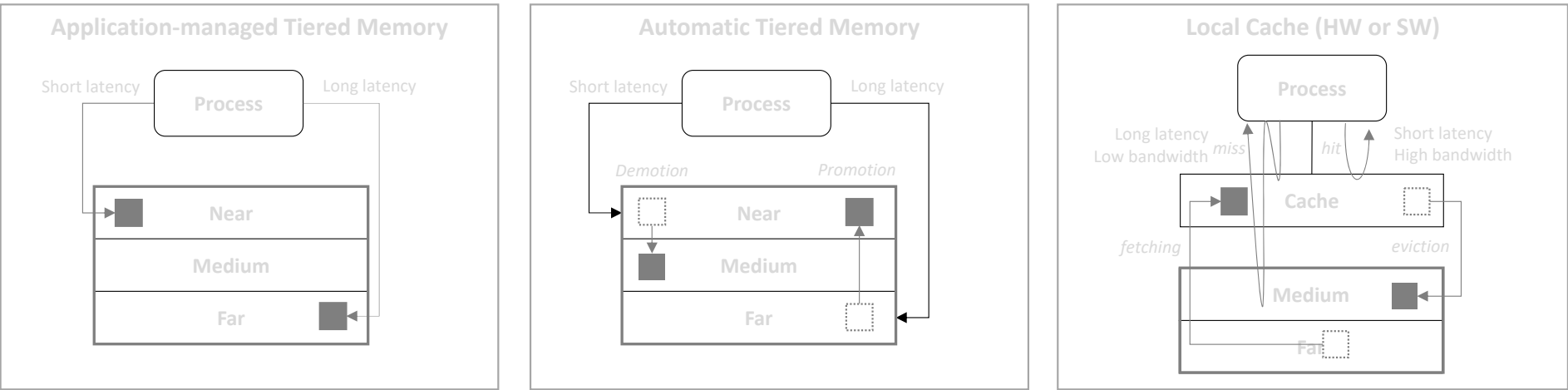
Memory Subsystem Architectures

- Toward disaggregated composable memory



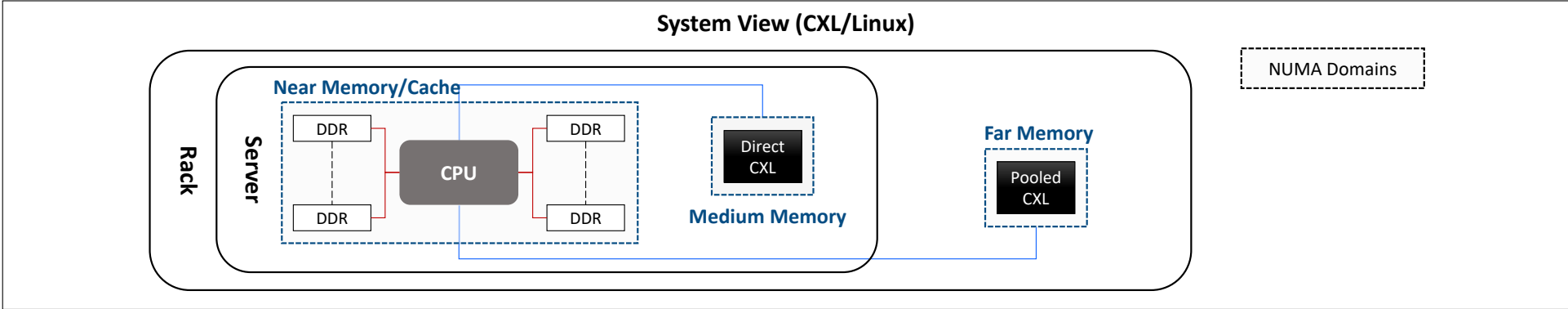
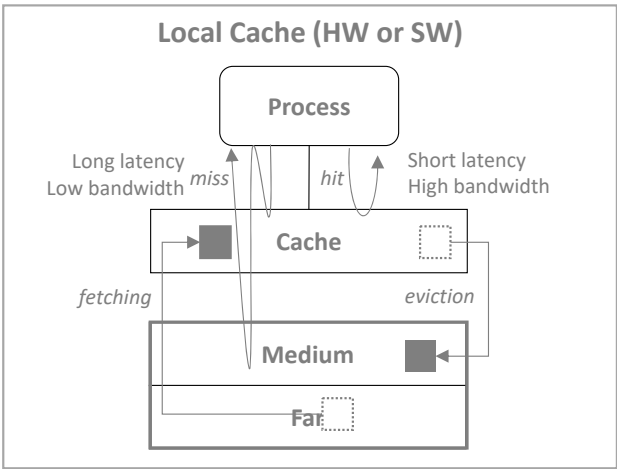
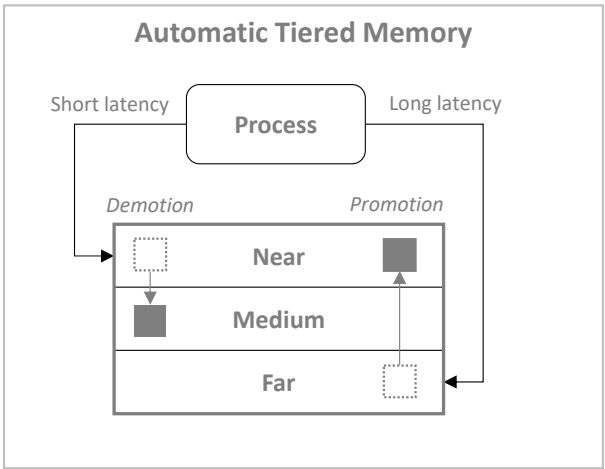
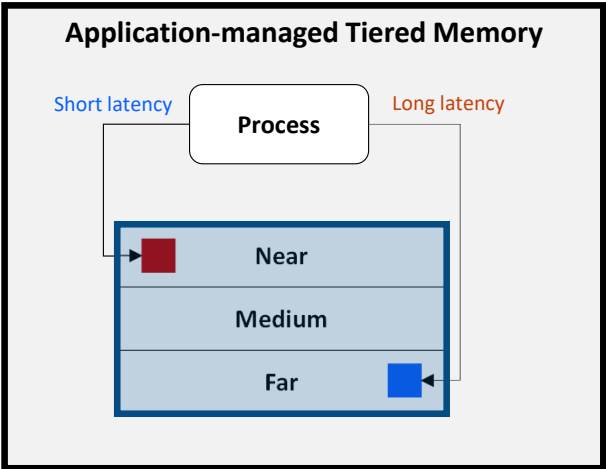
Hardware Abstraction for Memory Expander

- Multi-tiered memory management



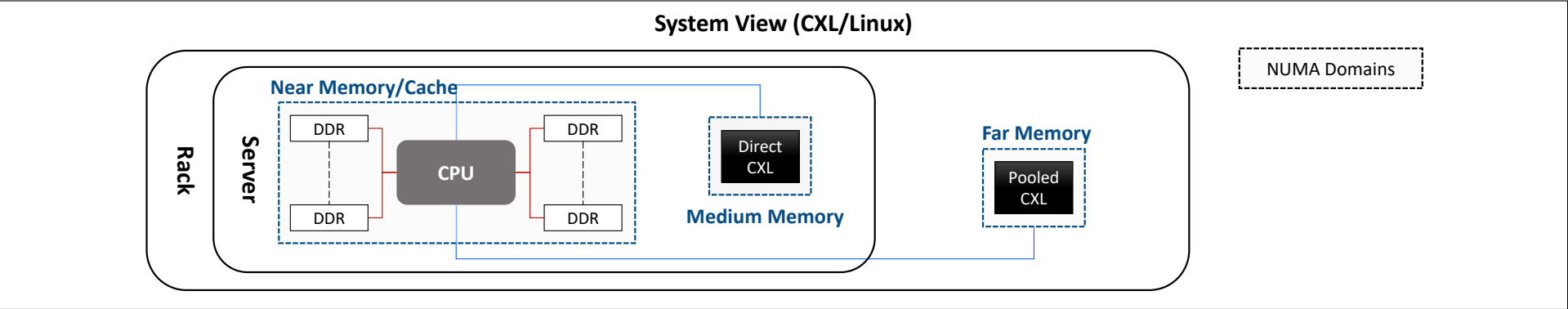
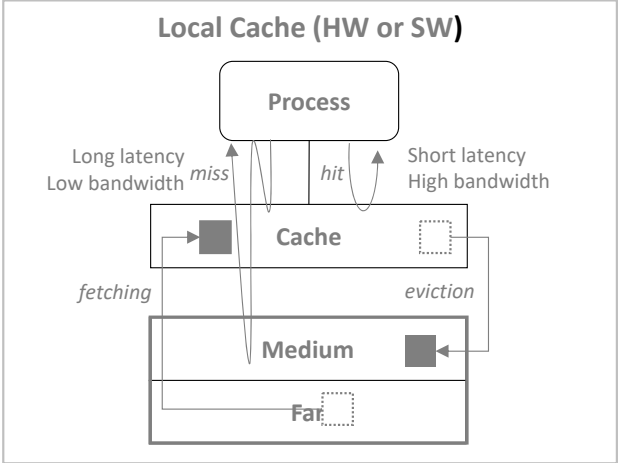
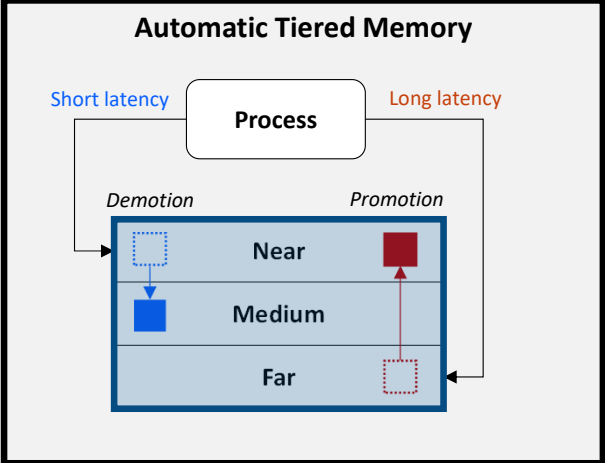
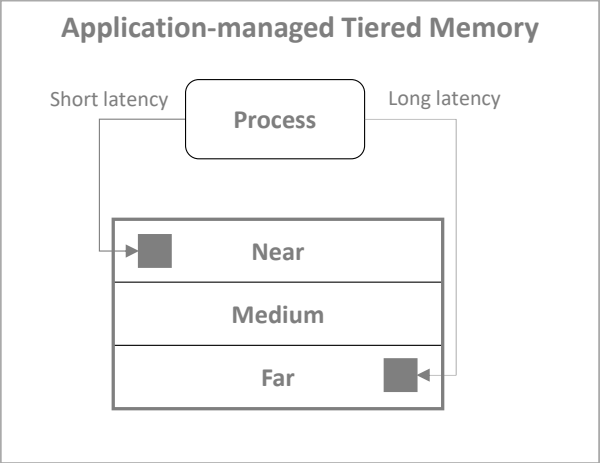
Basic Software for Memory Expander

- Multi-tiered memory management



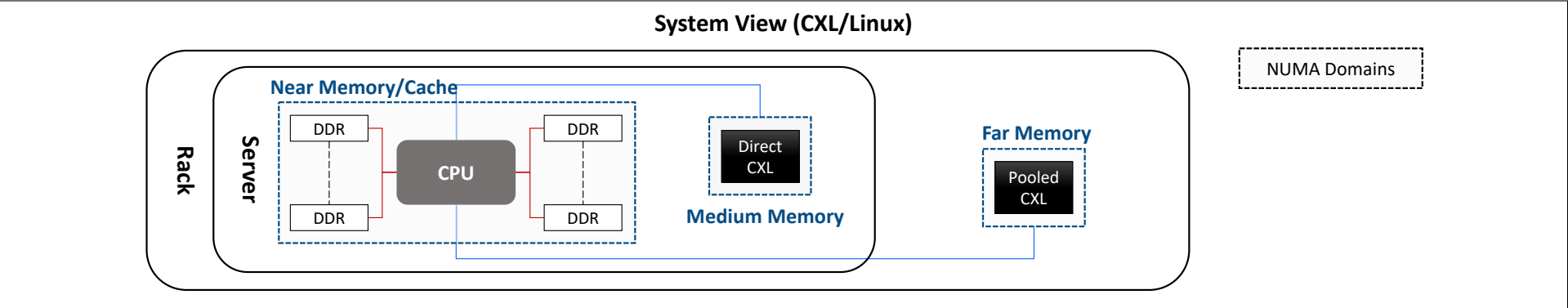
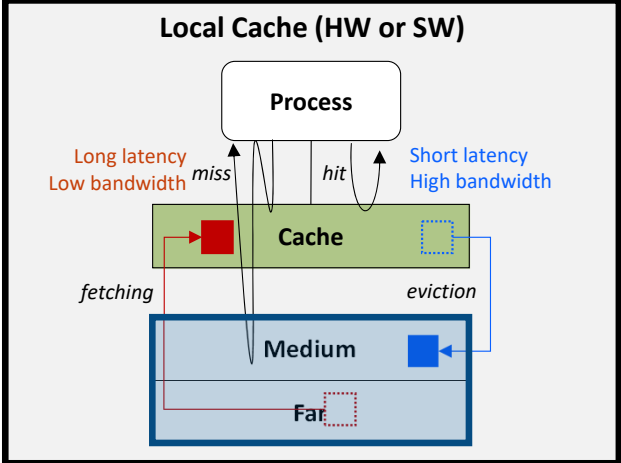
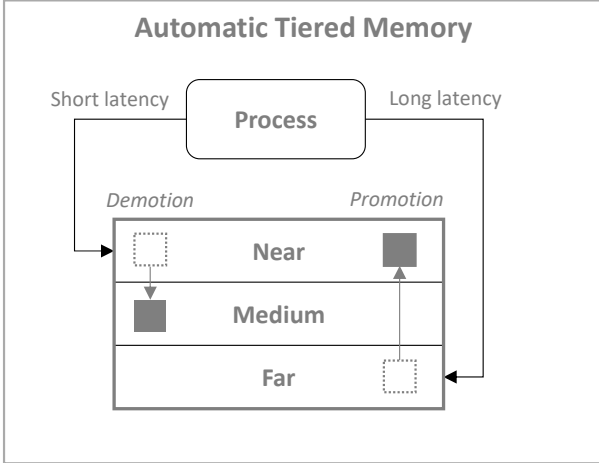
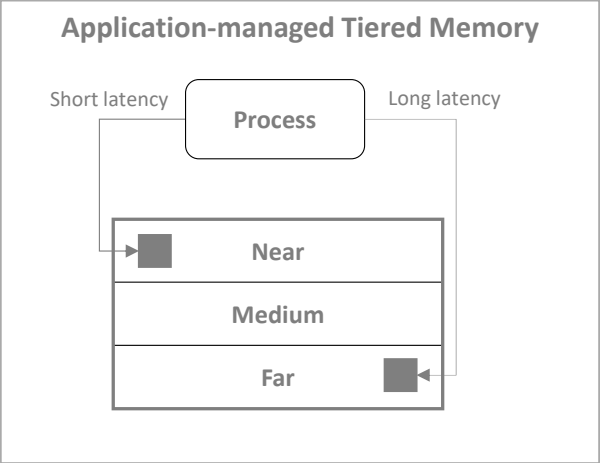
Basic Software for Memory Expander

- Multi-tiered memory management



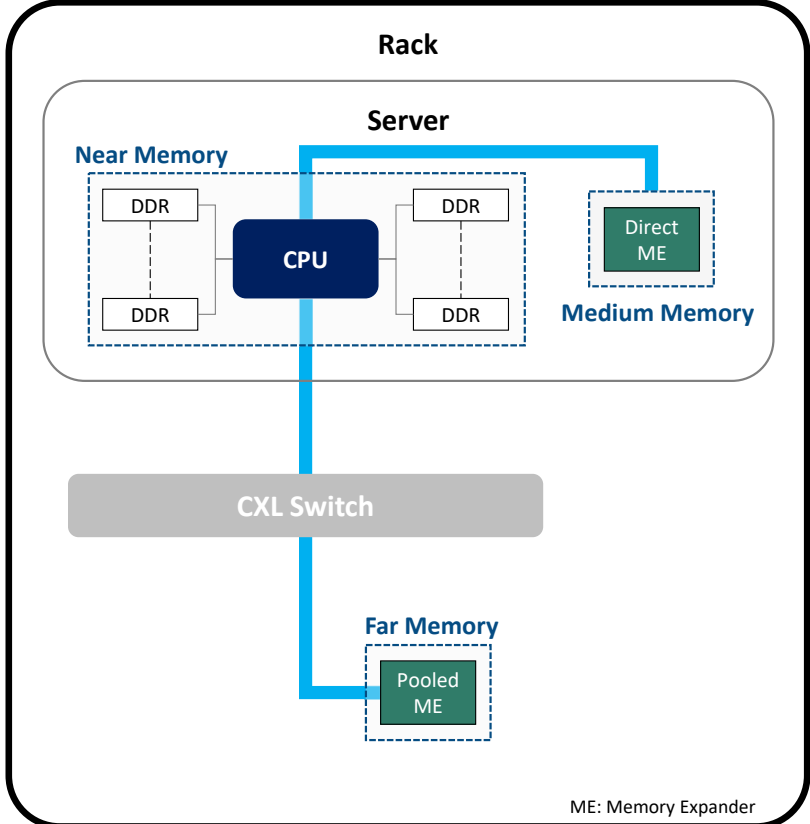
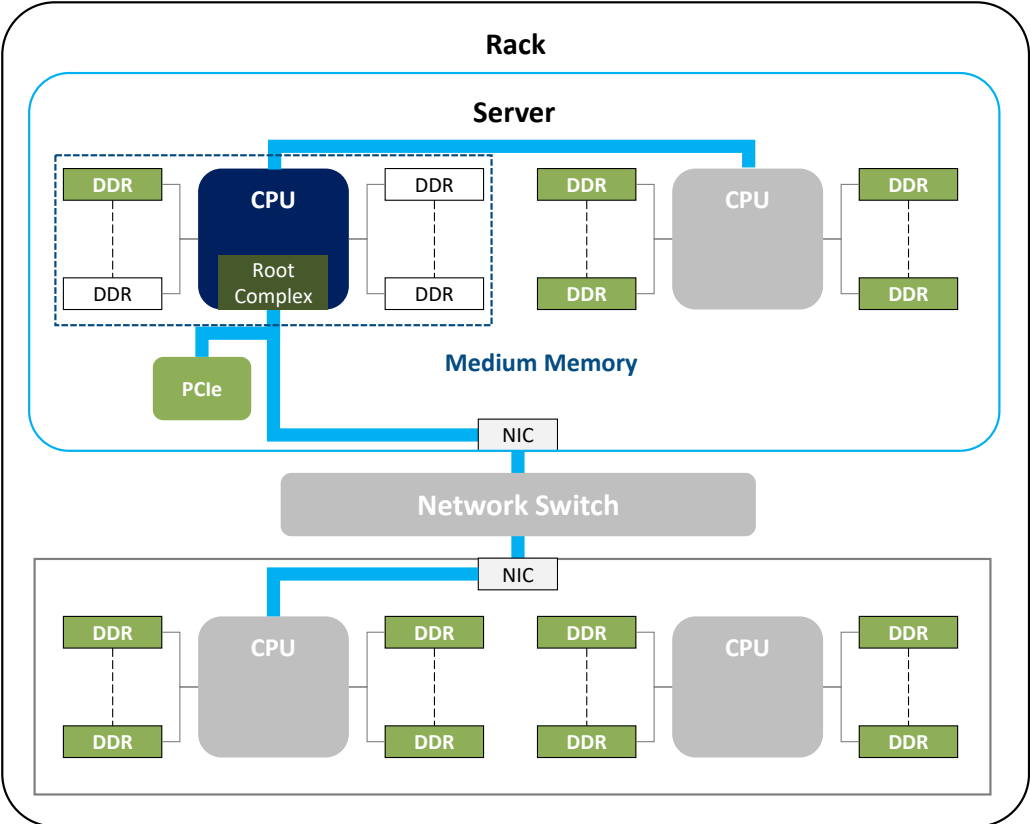
Basic Software for Memory Expander

- Multi-tiered memory management



Convergence of Memory Expansion Fabric into CXL

- Extended memory over DDR, processor interconnect, IO bus, and network into CXL-based memory



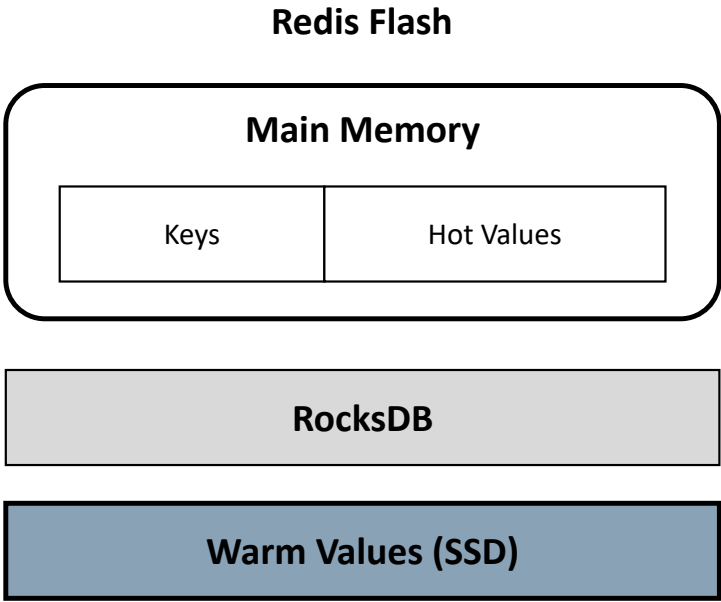
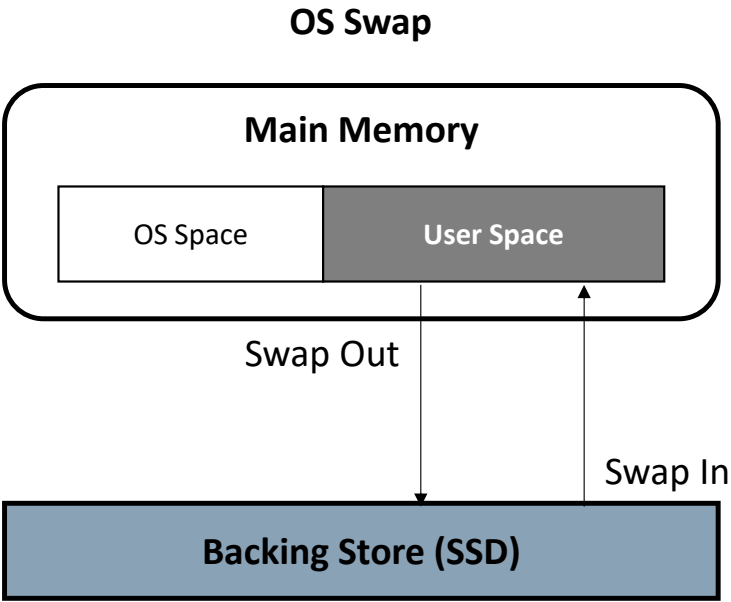
ME: Memory Expander



Convergence in Action: Storage Perspective

Economical Memory Challenge

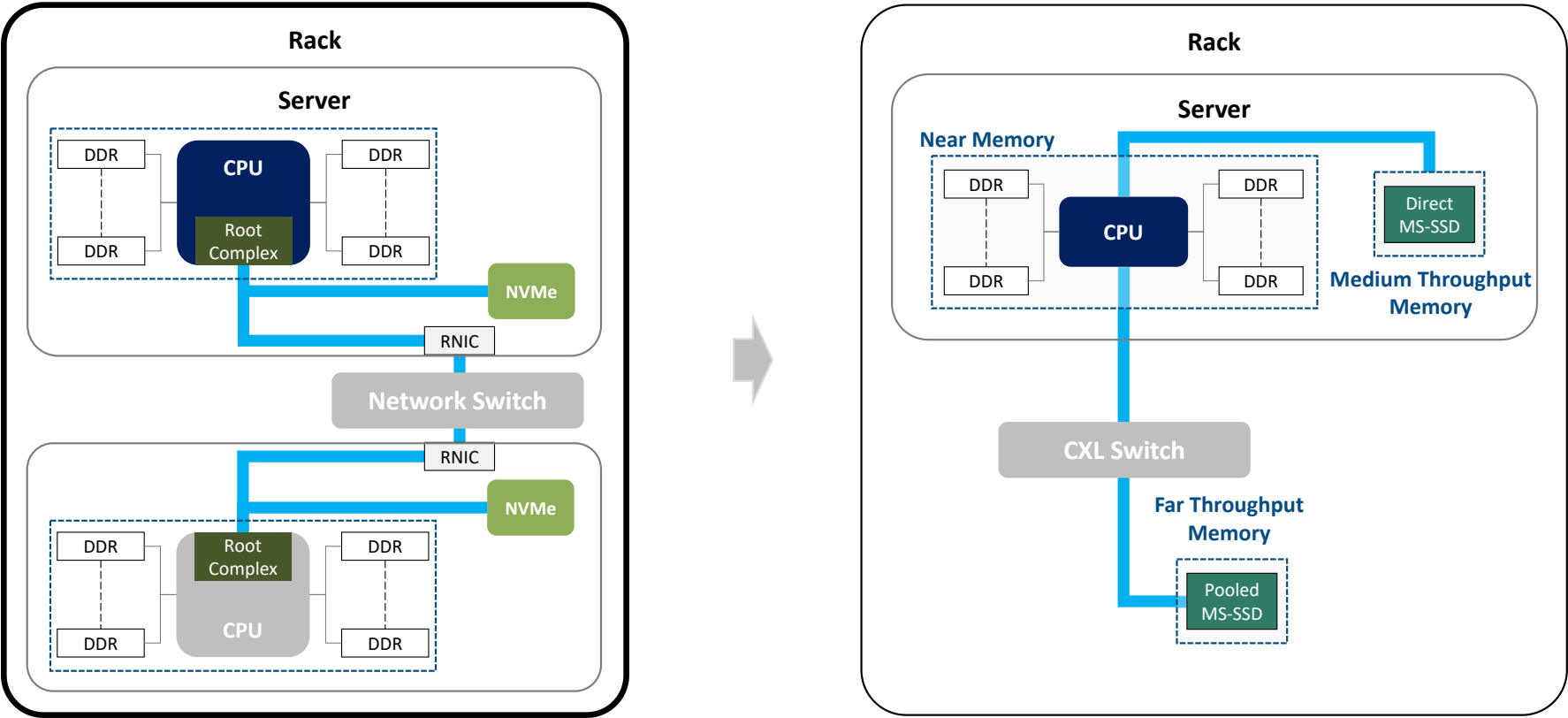
- Memory emulation using storage with large page movement through heavy software stack for better TCO



Large DRAM uneconomical for typical workloads

Convergence of Fast Small IOs into Memory Operations

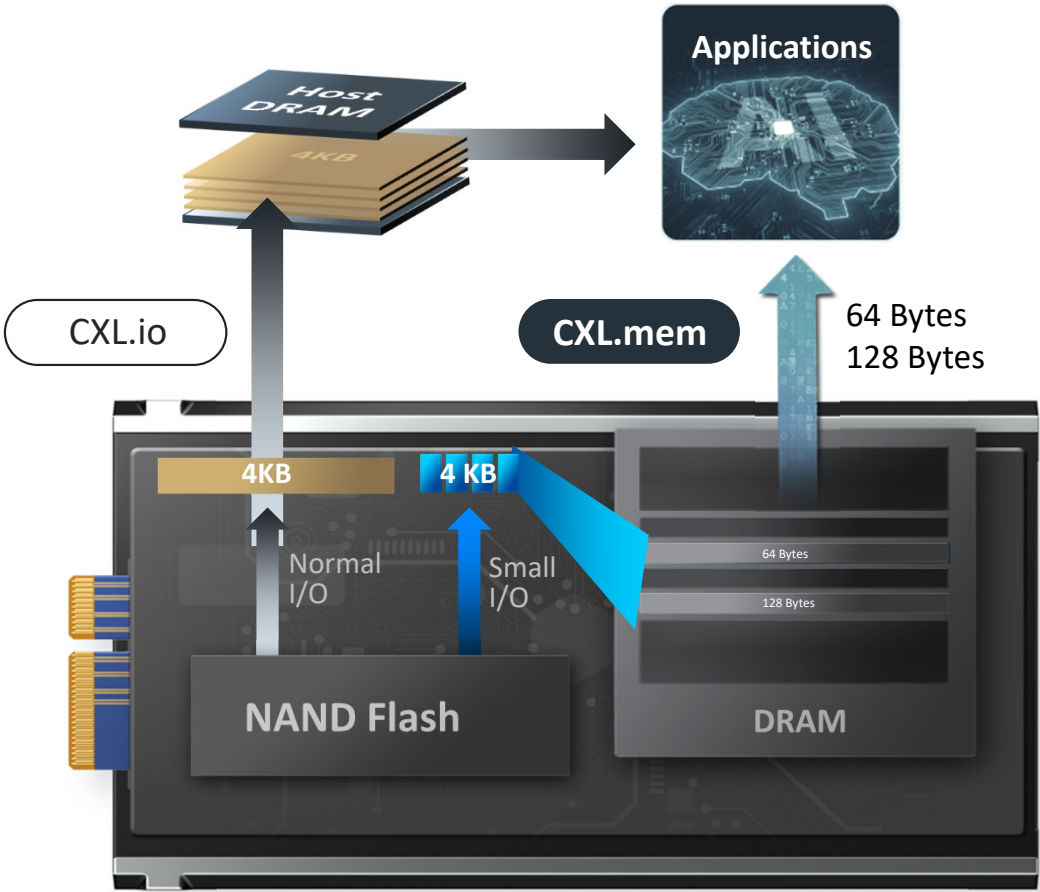
- Convergence of fast small IOs over DDR, processor interconnect, IO bus, and network into CXL-based memory operations



CXL-based SSD Concept

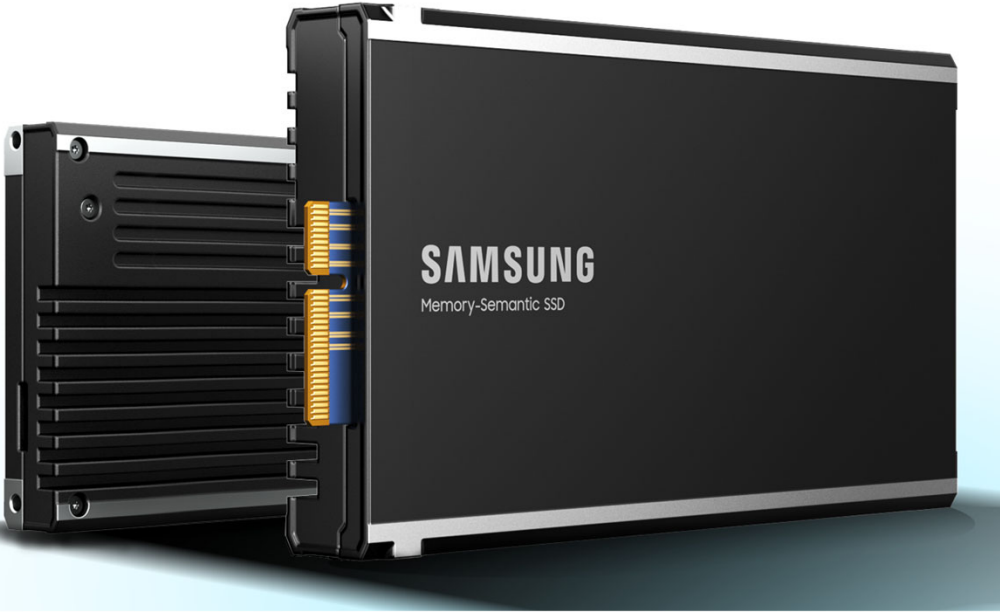
- Dual interface SSD with IO and memory for the same storage space

- Dual Mode Support
- Small Granularity Access
- Better System TCO



Samsung Memory Semantic SSD

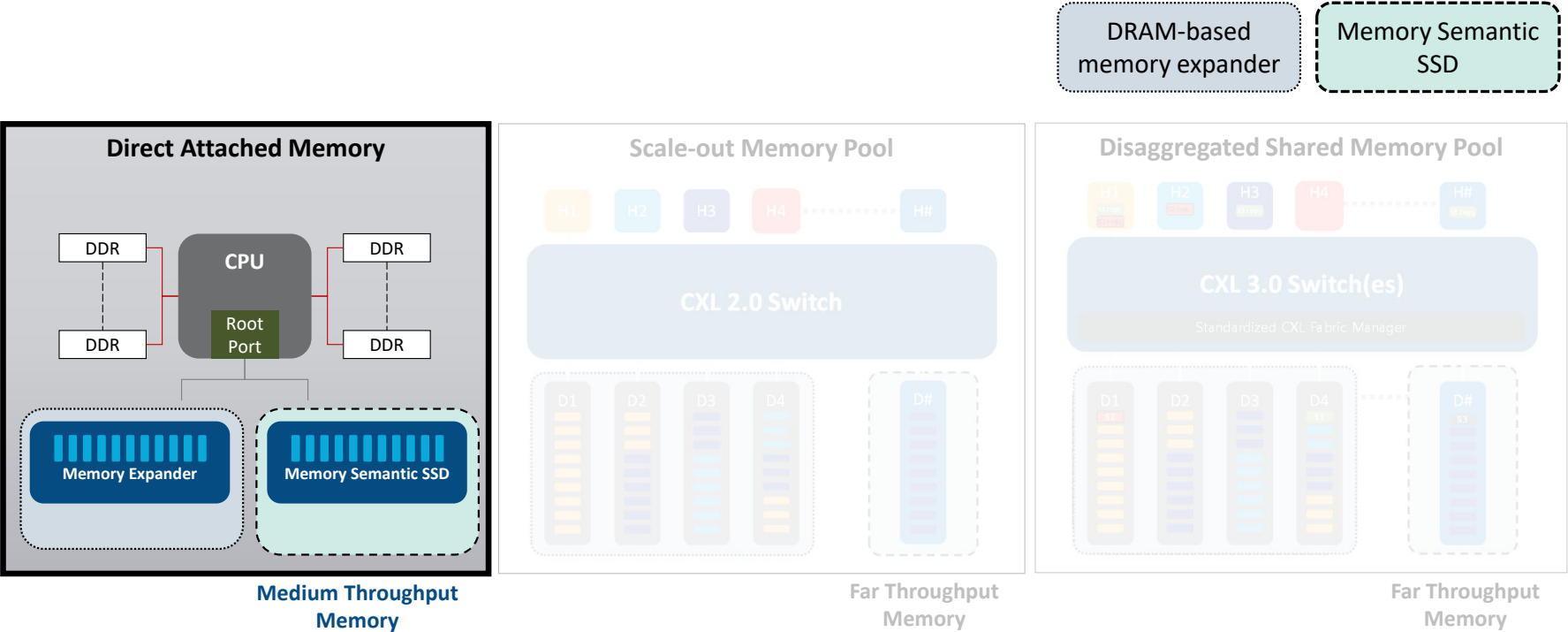
- Industry 1st CXL-based SSD concept PoC



Memory Semantic SSD	
Form Factor	E3.L
DRAM	16 GB DDR5
NAND	2TB
CXL Link	PCIe 4 X8
Max Bandwidth	16 GB/s
IOPS	Max 20MIOPS
Latency	< 1us
Version	CXL 2.0
Device Type	Type-2
RAS	TBD
Availability	TBD

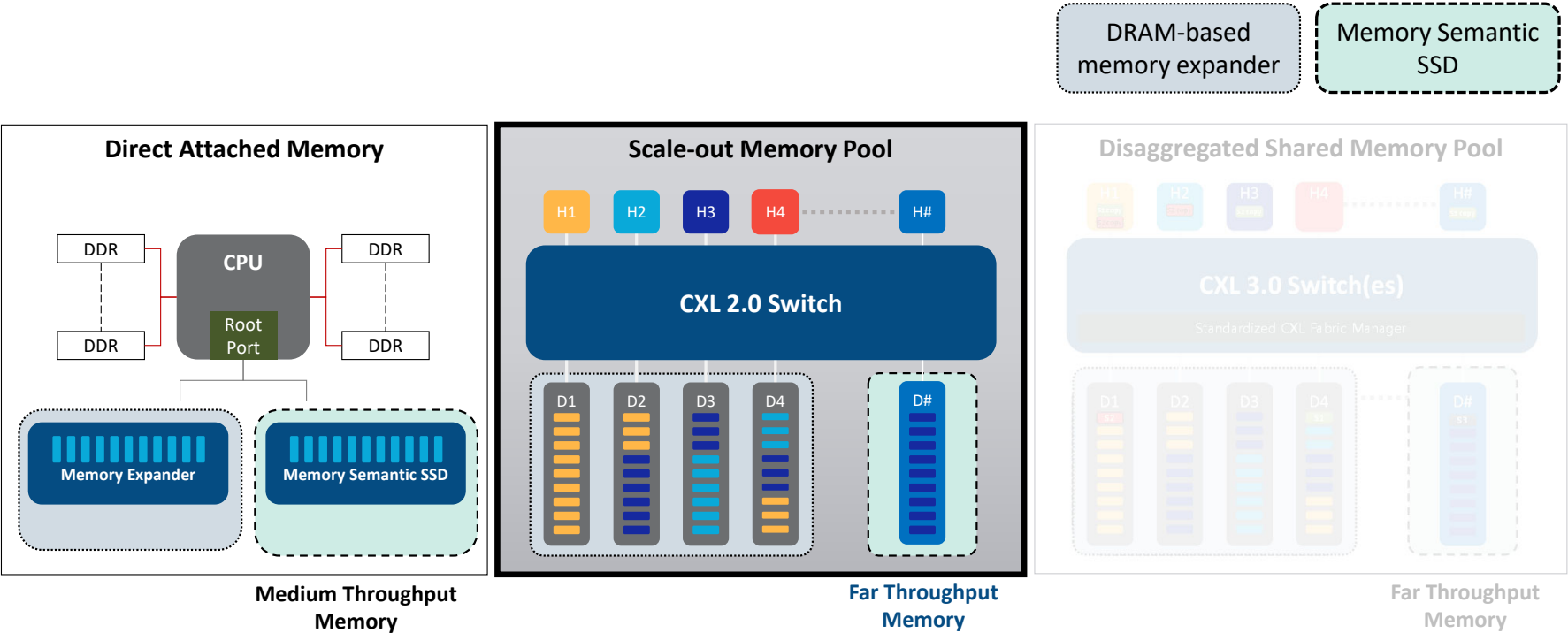
Extended Memory Subsystem Architectures

- Toward disaggregated composable memory with multi-tiered memory



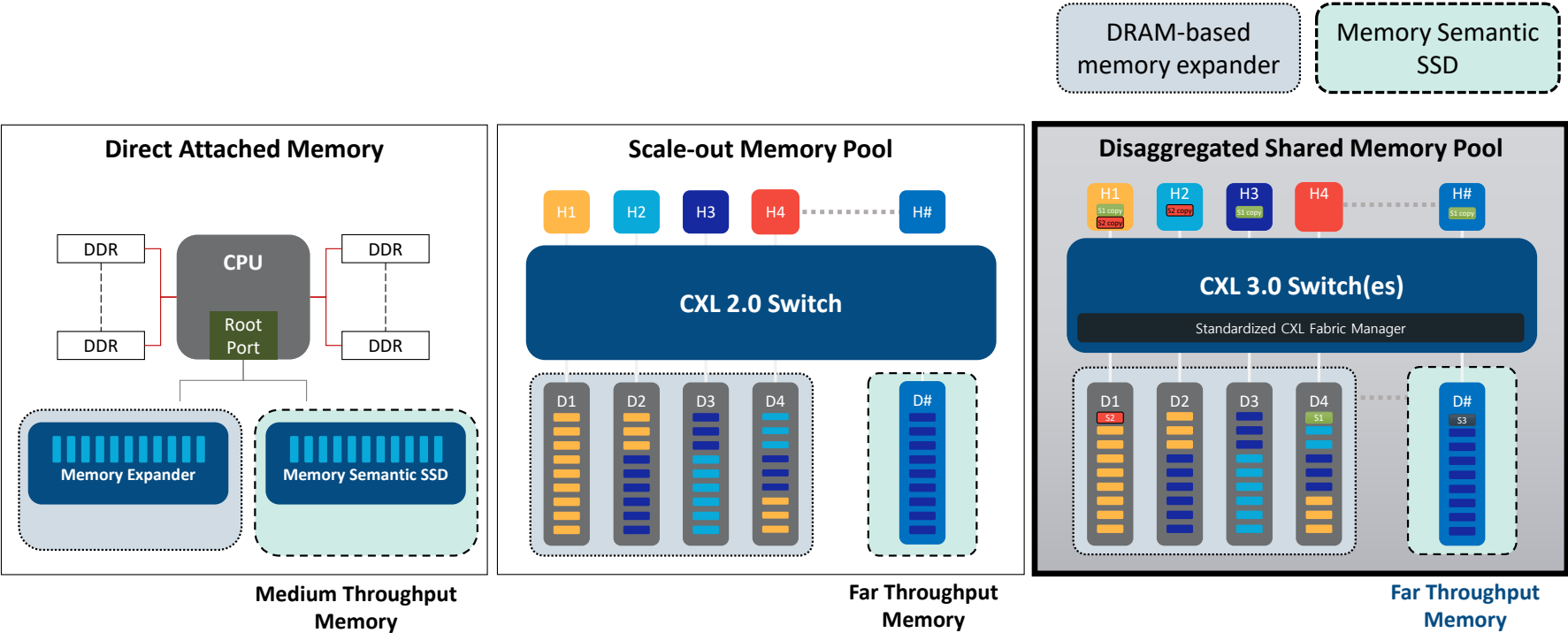
Extended Memory Subsystem Architectures

- Toward disaggregated composable memory with multi-tiered memory



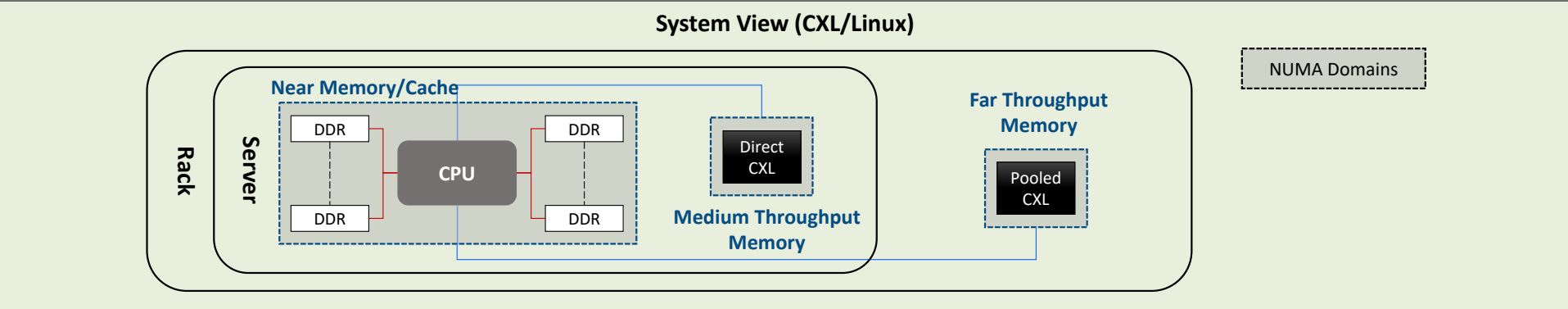
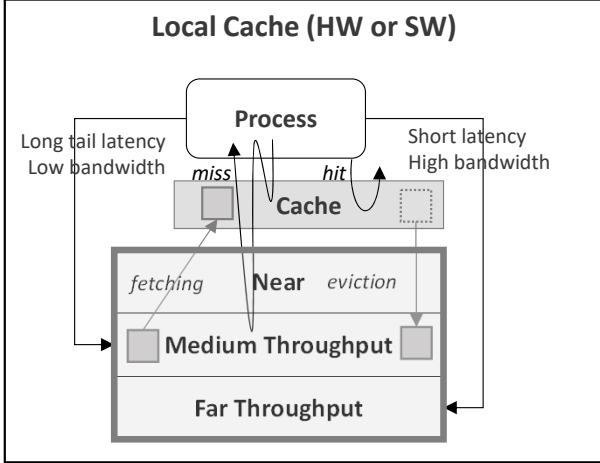
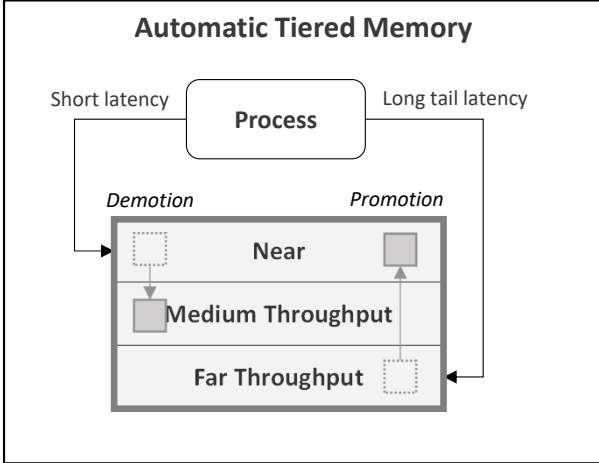
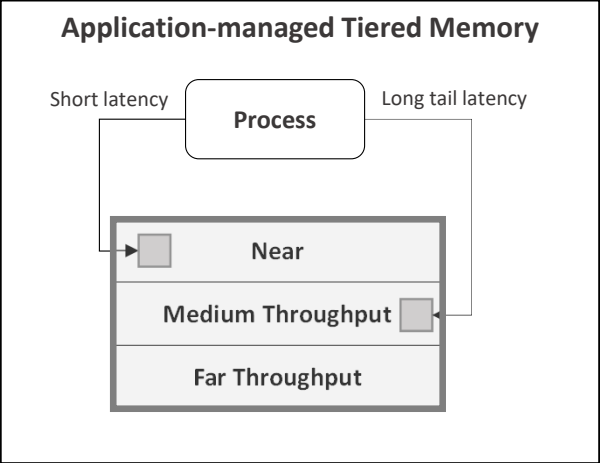
Extended Memory Subsystem Architectures

- Toward disaggregated composable memory with multi-tiered memory



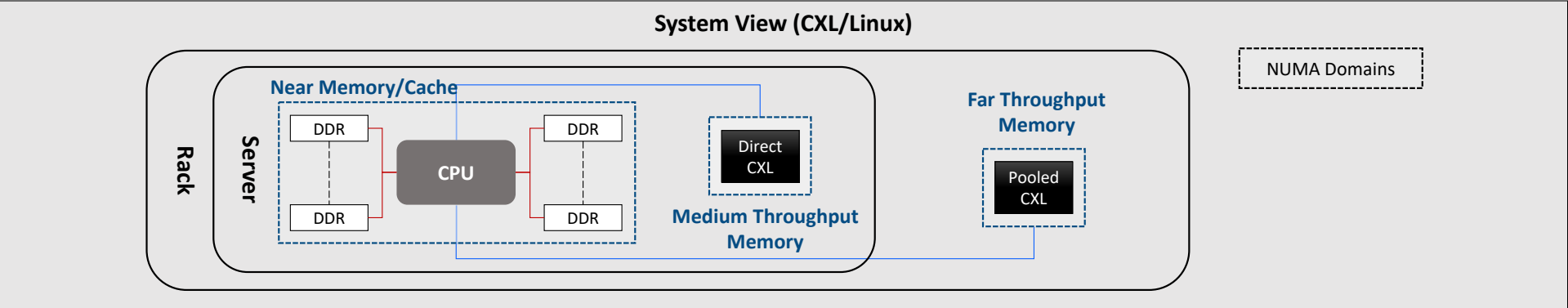
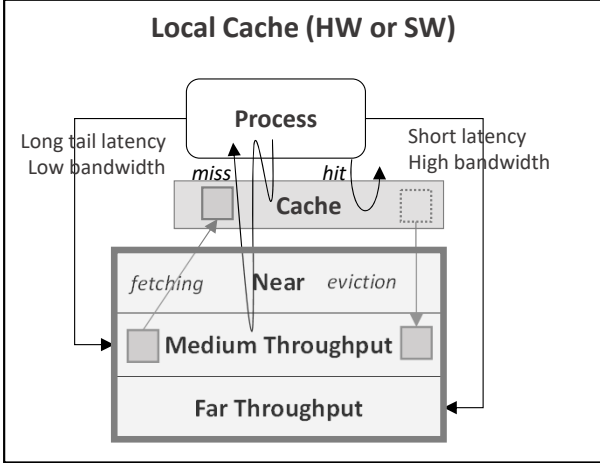
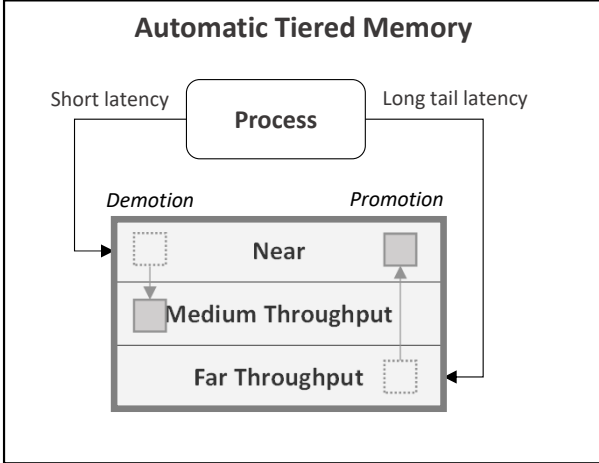
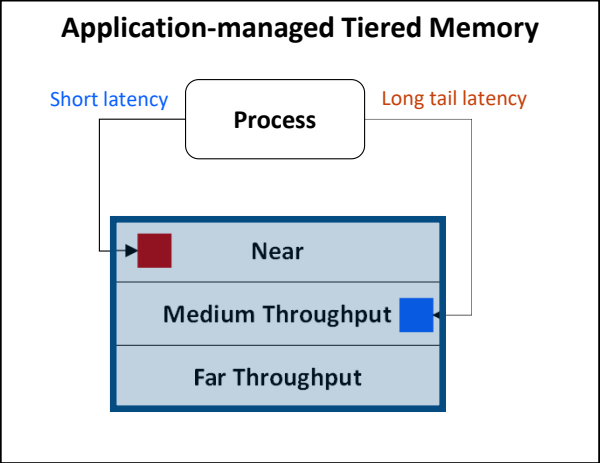
Hardware Abstraction for Memory Semantic SSD

- Tiered memory management for long tail latency memory



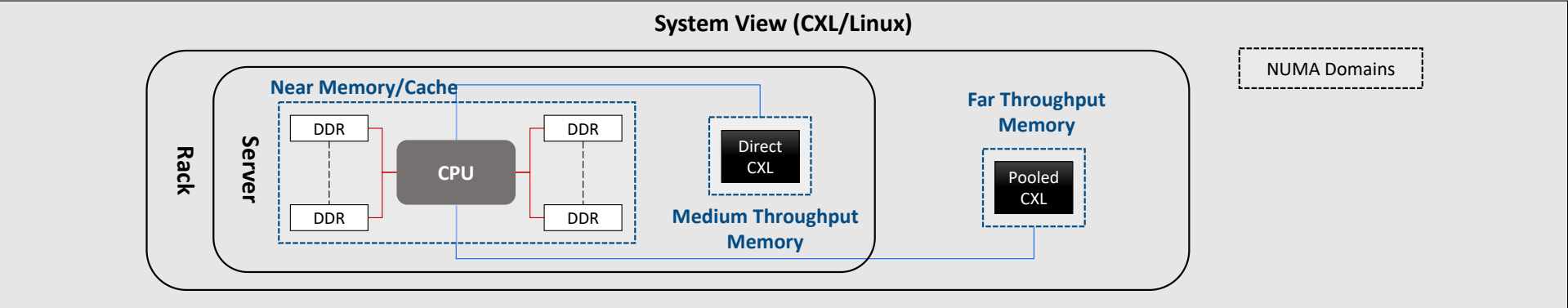
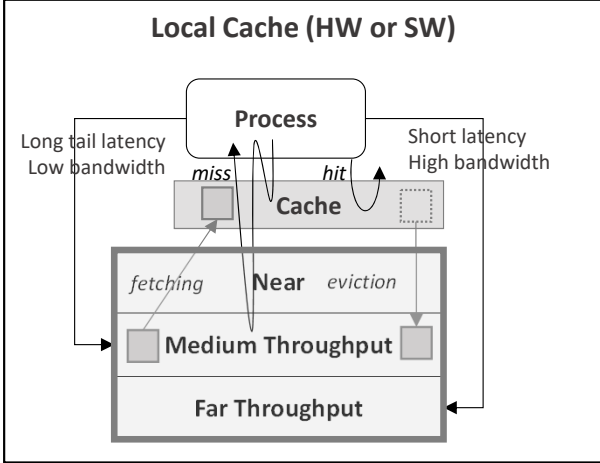
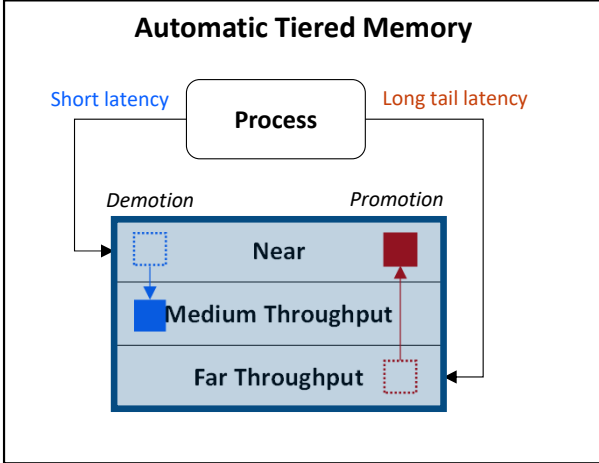
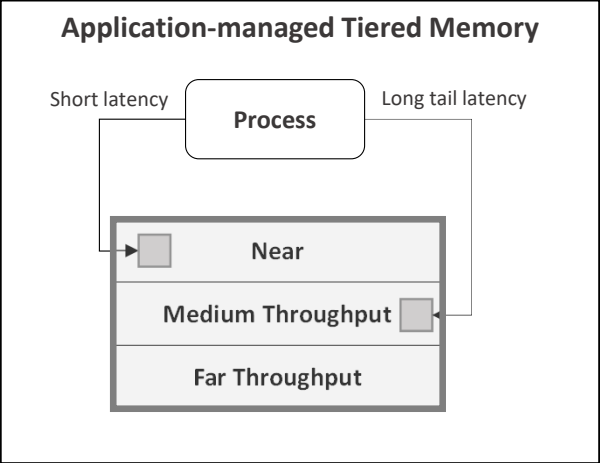
Basic Software for Memory Semantic SSD

- Tiered memory management for long tail latency memory



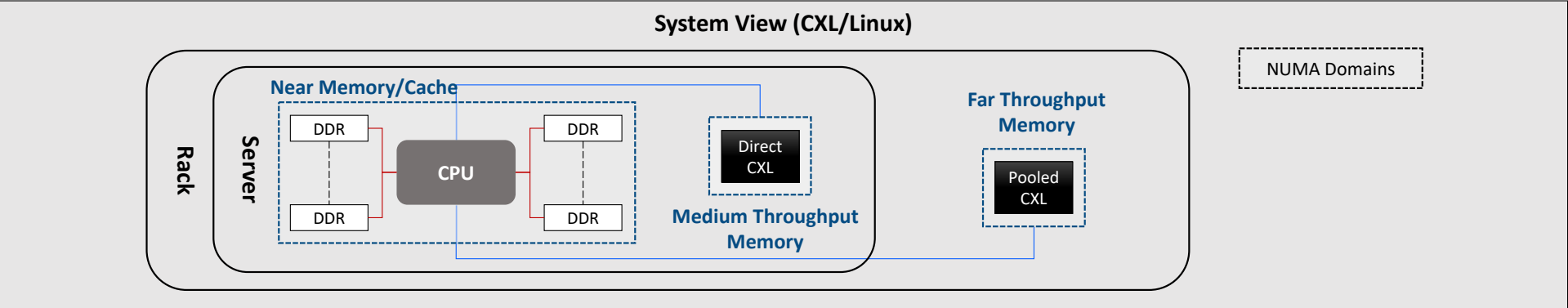
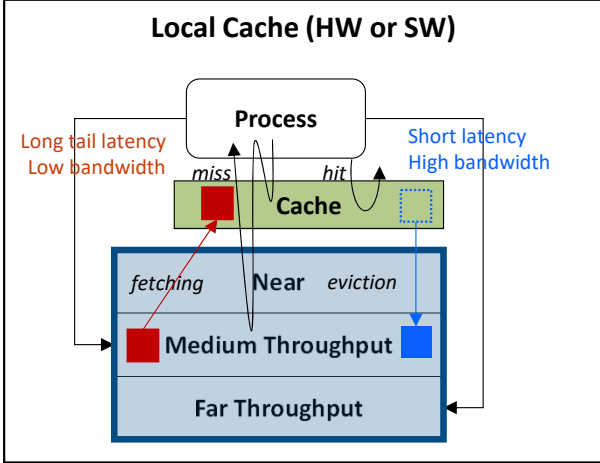
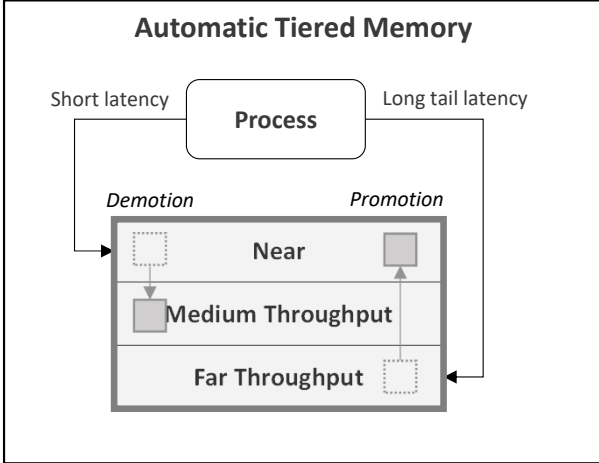
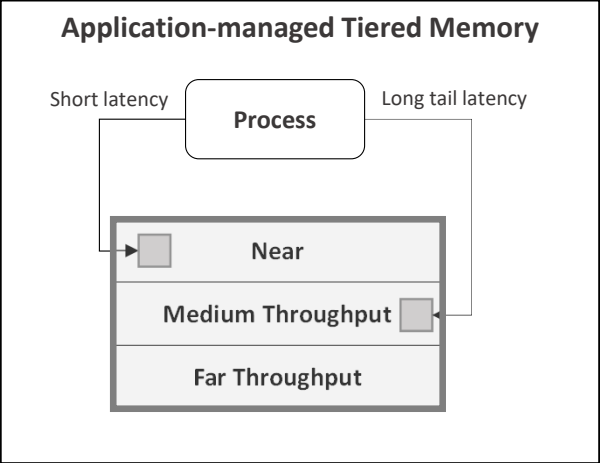
Basic Software for Memory Semantic SSD

- Tiered memory management for long tail latency memory



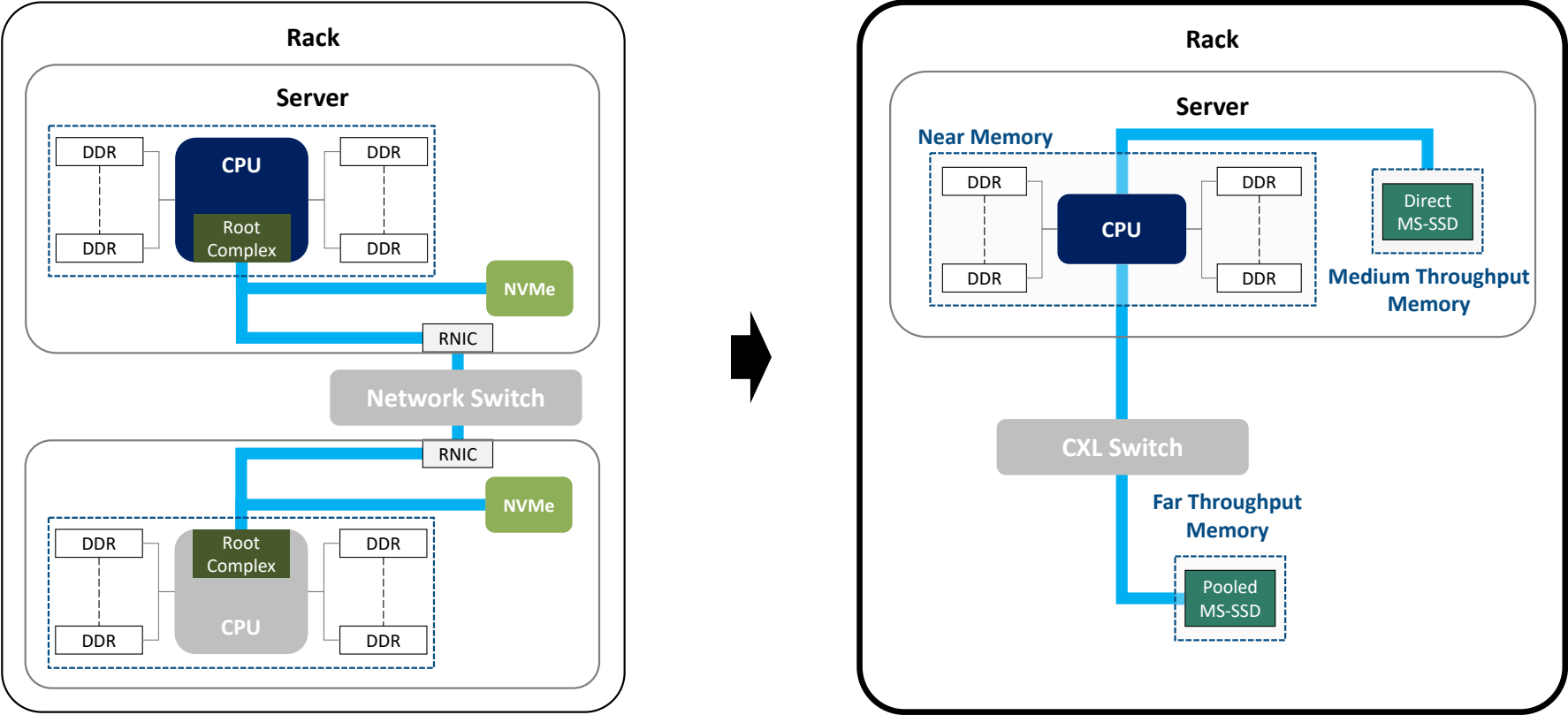
Basic Software for Memory Semantic SSD

- Tiered memory management for long tail latency memory



Convergence of Fast Small IOs into Memory Operations

- Convergence of fast small IOs over DDR, processor interconnect, IO bus, and network into CXL-based memory operations

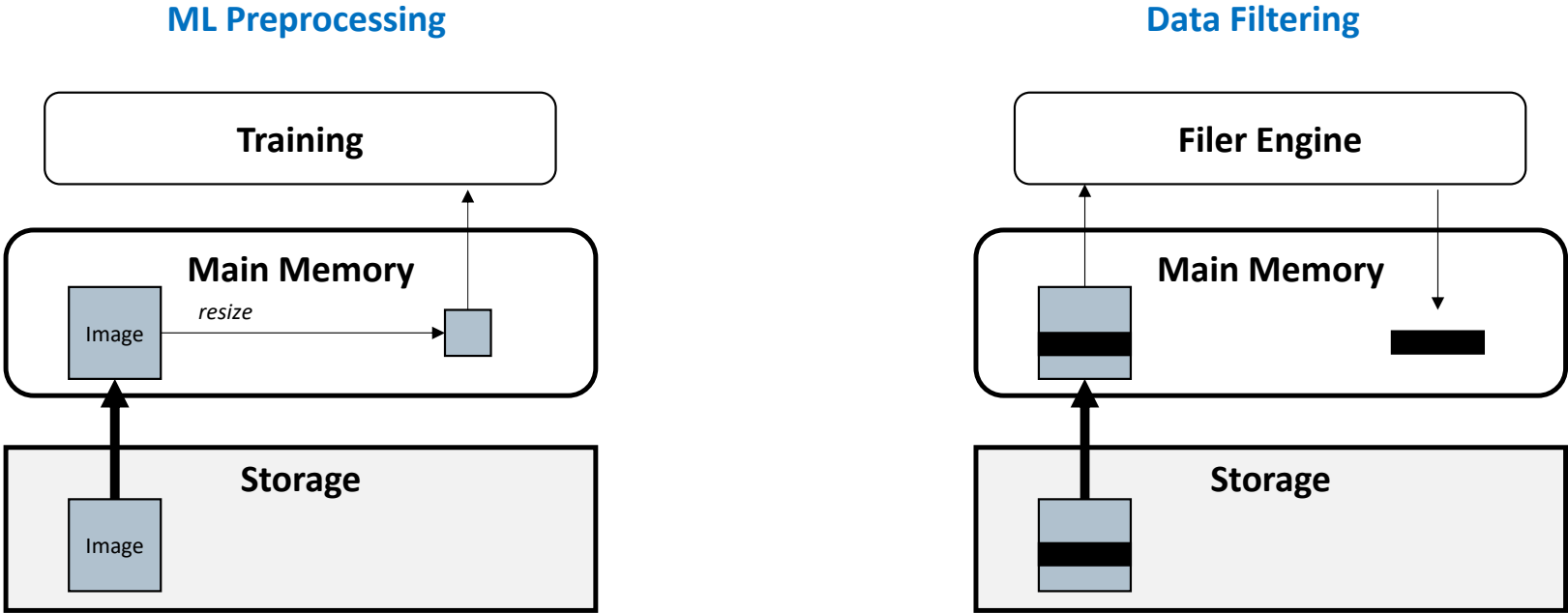




Convergence in Action: Computing Perspective

CPU-centric Computing Challenge

- Inefficient use of bandwidth and power across memory, storage, buses, network, and CPU

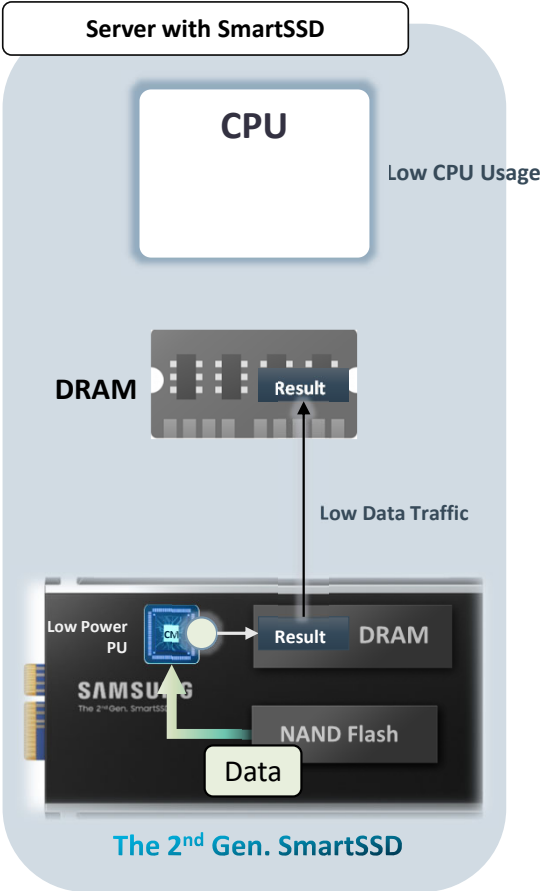
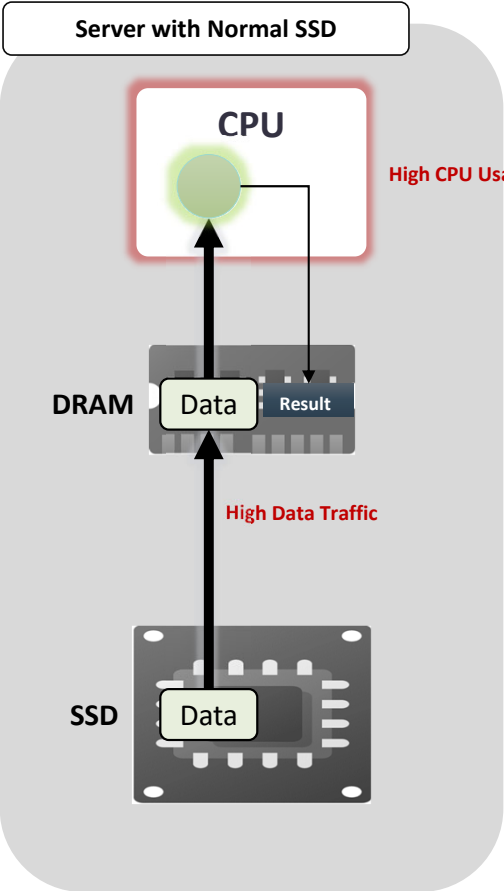


Page-based resource management across memory and storage

Computational Storage/Memory Device Concept

- Power-efficient near data processing in storage/memory devices

- Low Power
- Data Reduction
- High Effective BW



Samsung 2nd Generation SmartSSD

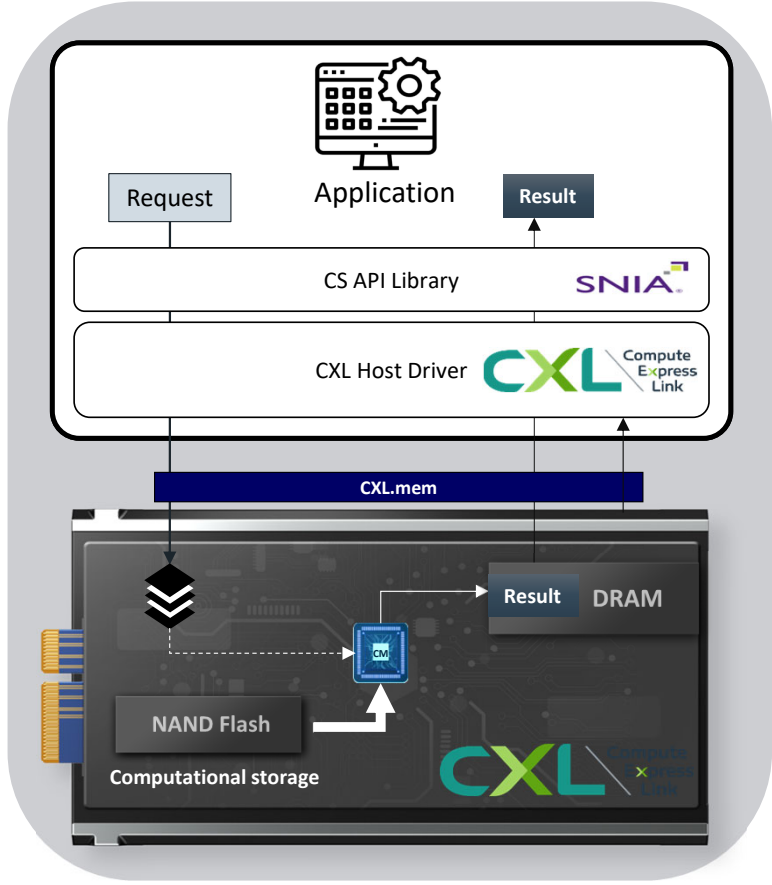
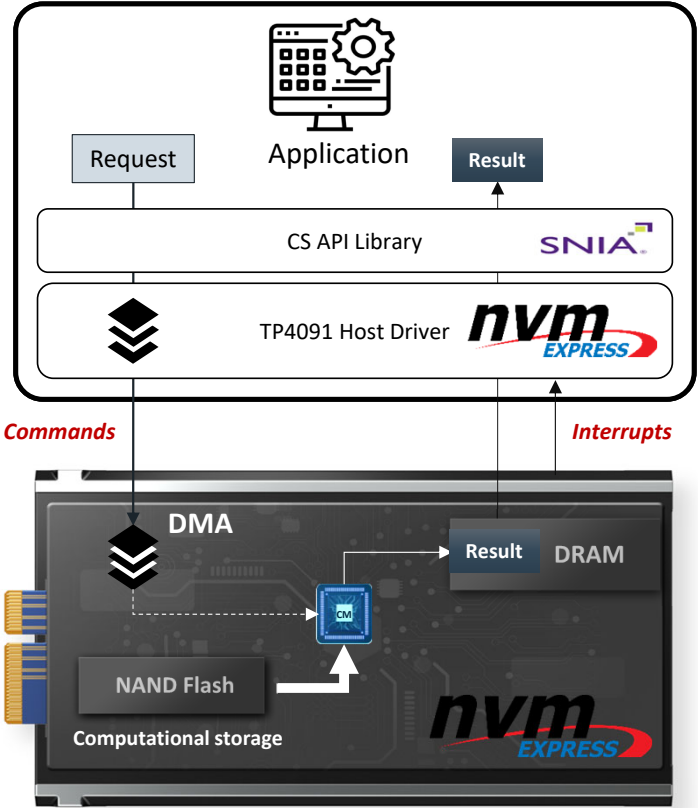
- NVMe TP4091-compatible computational storage device PoC



Samsung 2 nd Gen SmartSSD	
Form Factor	E3.L
DRAM	16 GB DDR5
NAND	2TB
NVMe Link	PCIe 4 X4
Max Bandwidth	8 GB/s
IOPS	1 MIOPS
Protocol	NVMe TP4091
API	SNIA CS API

Convergence of Computing into Memory Abstraction

- Control/data planes for computational memory/storage can be converted into memory operations

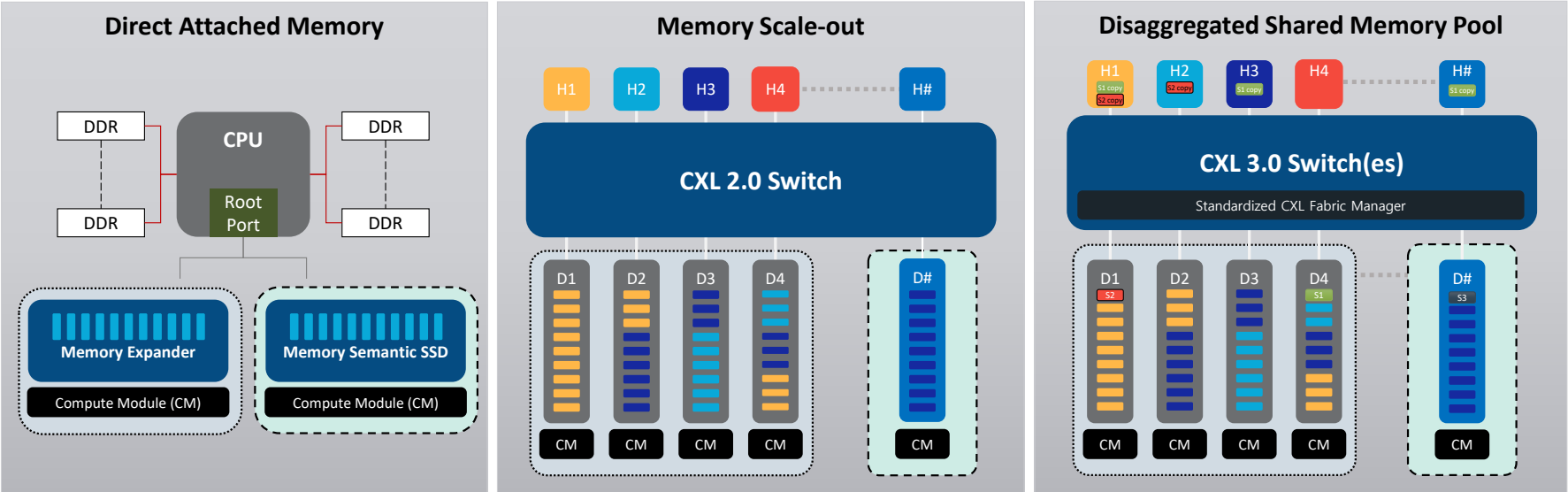
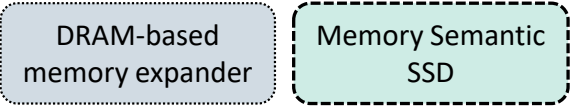




Memory Technology Continuum

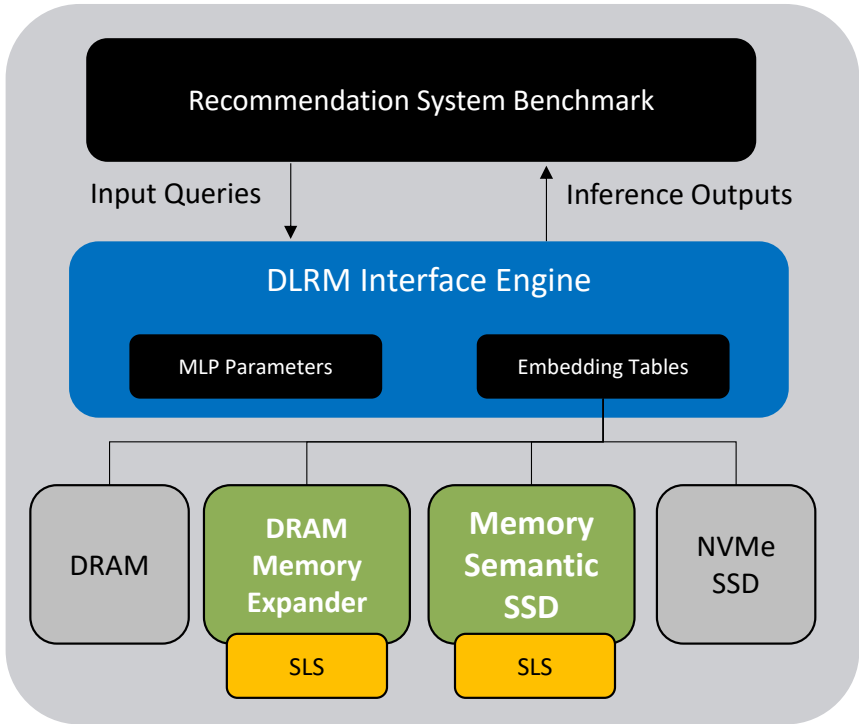
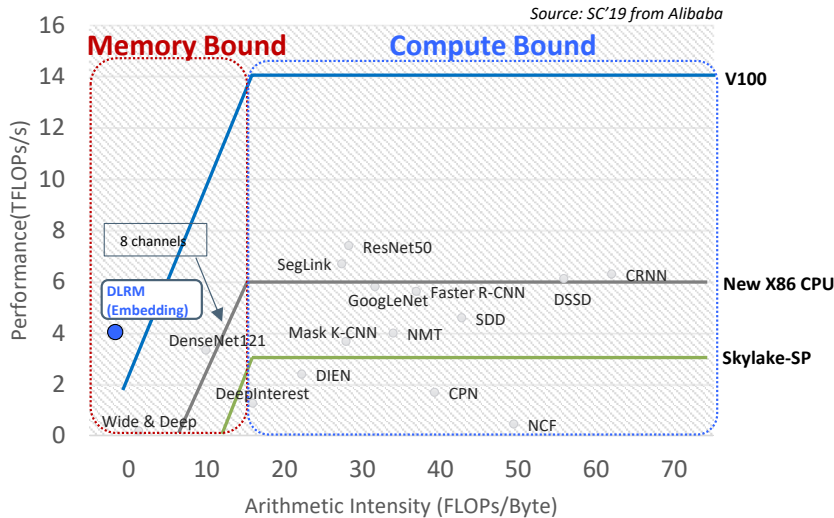
Converged Memory Subsystem Architectures

- Memory, storage, and computing functions are converged through consistent memory interfaces enabled by CXL



Convergence of Memory, Storage, and Acceleration

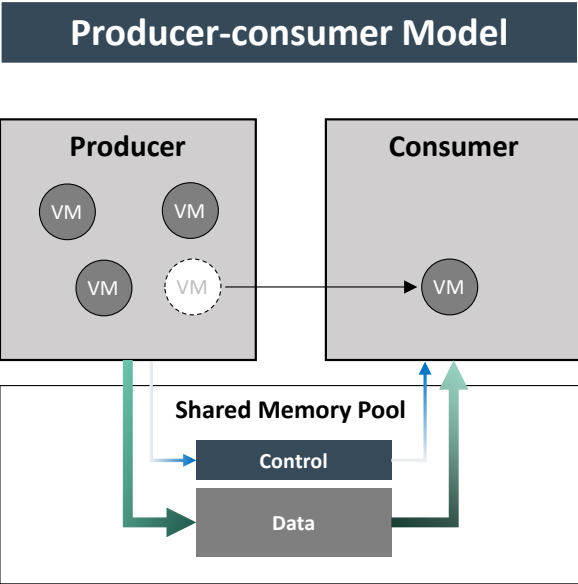
- Memory Expanders/Memory Semantic SSDs can provide additional bandwidth, capacity, and throughput for DLRM



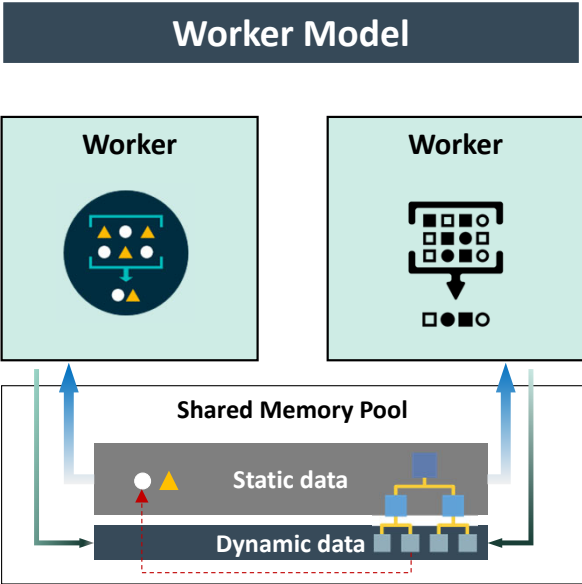
SLS: Sparse Length Sum

Convergence of Network, Memory, and Storage

- Message passing over network can be replaced with global shared memory over CXL



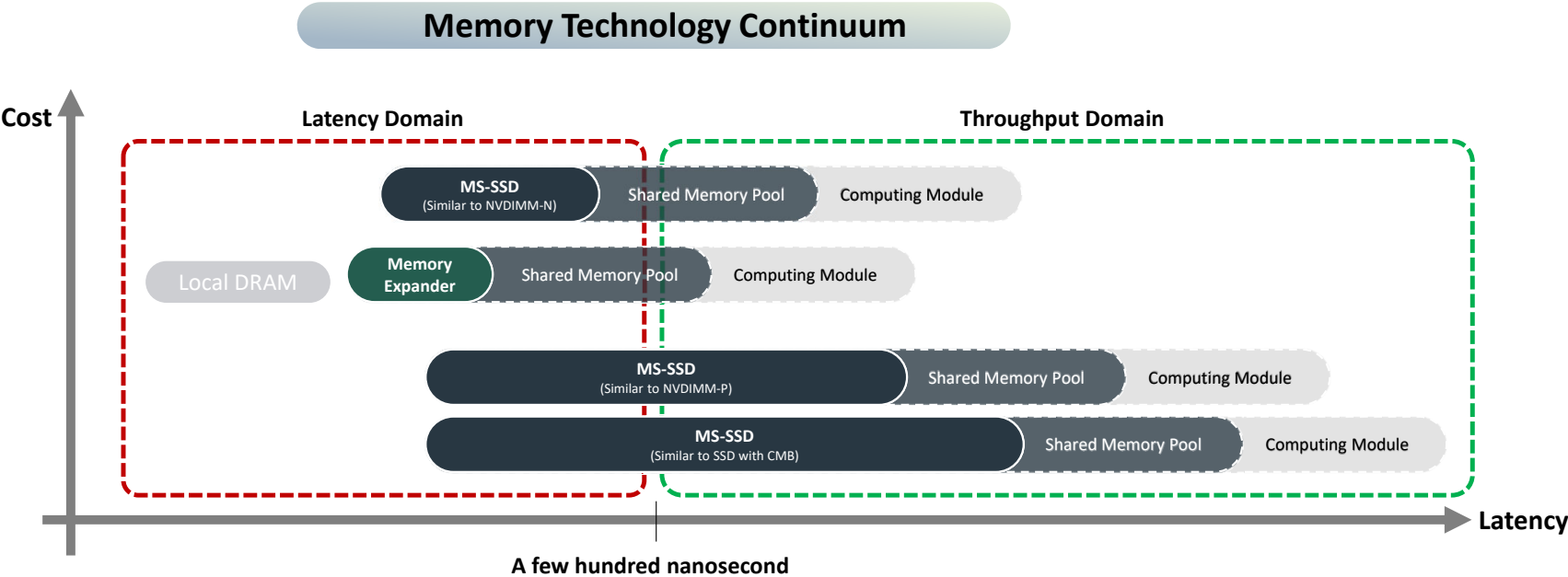
- Example*
- VM migration



- Example*
- Deduplication finger print
 - In-memory database main store

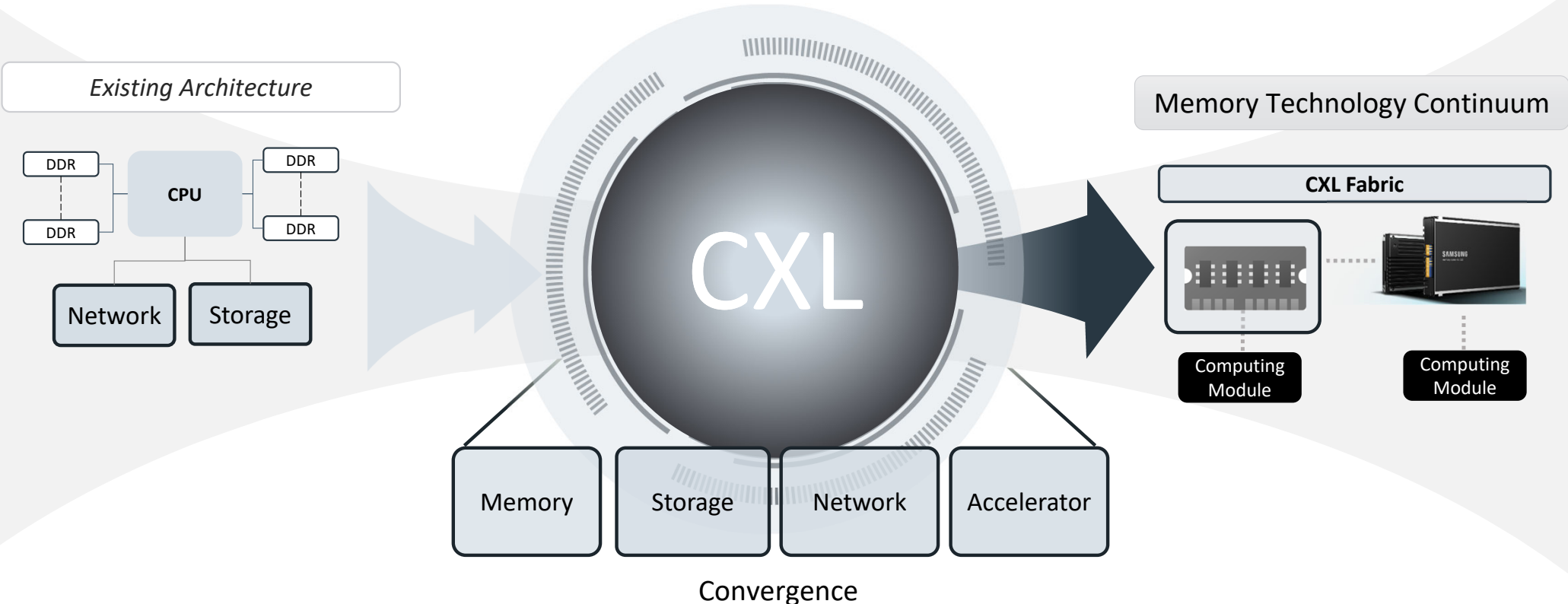
Memory Technology Continuum

- The application constructs the best option in the memory technology continuum as needed.



Conclusion

Why Should the Storage Community Care about CXL?





Thank You!

www.smartssd.samsung.com <https://samsungmsl.com/>