COMPUTE + MEMORY + STORAGE SUMMIT

Architectures, Solutions, and Community
VIRTUAL EVENT, APRIL 11-12, 2023

Compute Express Link™ (CXL™) 3.0: Expanded capabilities for increasing scale and optimizing resource utilization

Andy Rudoff, Intel

Compute Express Link (CXL)

CXL Overview

- New breakthrough high-speed fabric
  - Enables a high-speed, efficient interconnect between CPU, memory and accelerators
  - Builds upon PCI Express® (PCIe®) infrastructure, leveraging the PCIe® physical and electrical interface
  - Maintains memory coherency between the CPU memory space and memory on CXL attached devices
    - Enables fine-grained resource sharing for higher performance in heterogeneous compute environments
    - Enables memory disaggregation, memory pooling and sharing, persistent memory and emerging memory media
- Delivered as an open industry standard
  - CXL 3.0 specification is fully backward compatible with CXL 2.0 and CXL 1.1
  - Future CXL Specification generations will include continuous innovation to meet industry needs and support new technologies

CXL Board of Directors

Industry Open Standard for High Speed Communications

200+ Member Companies

Confidential | CXL™ Consortium 2022

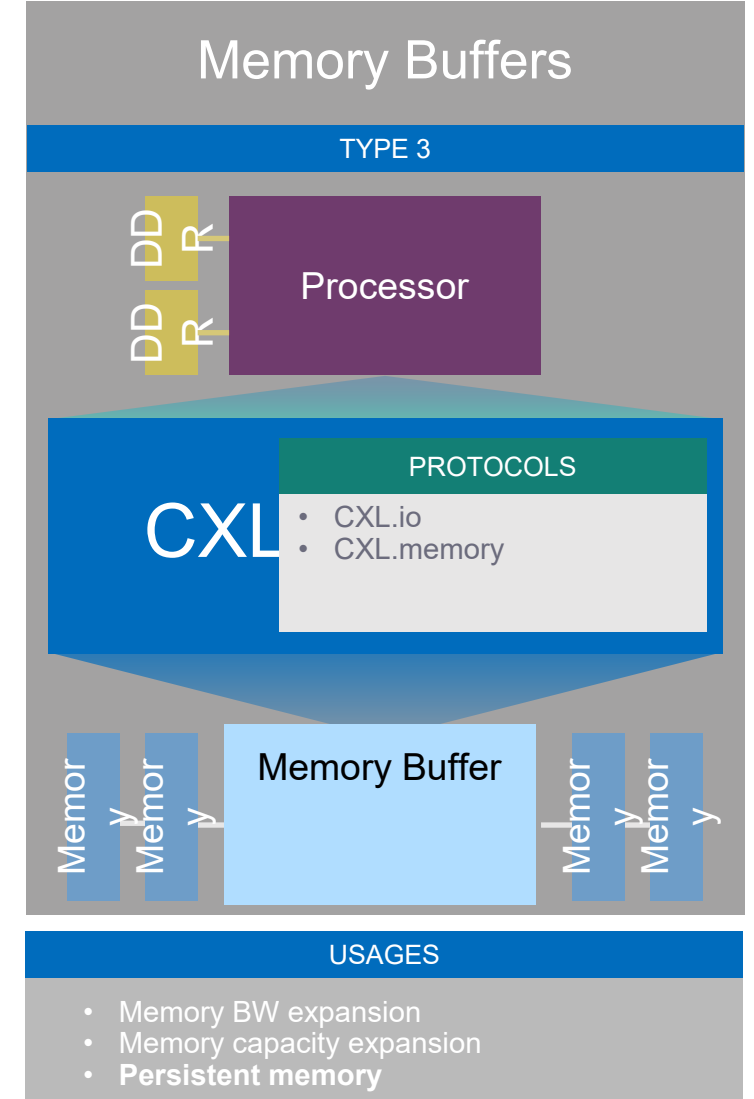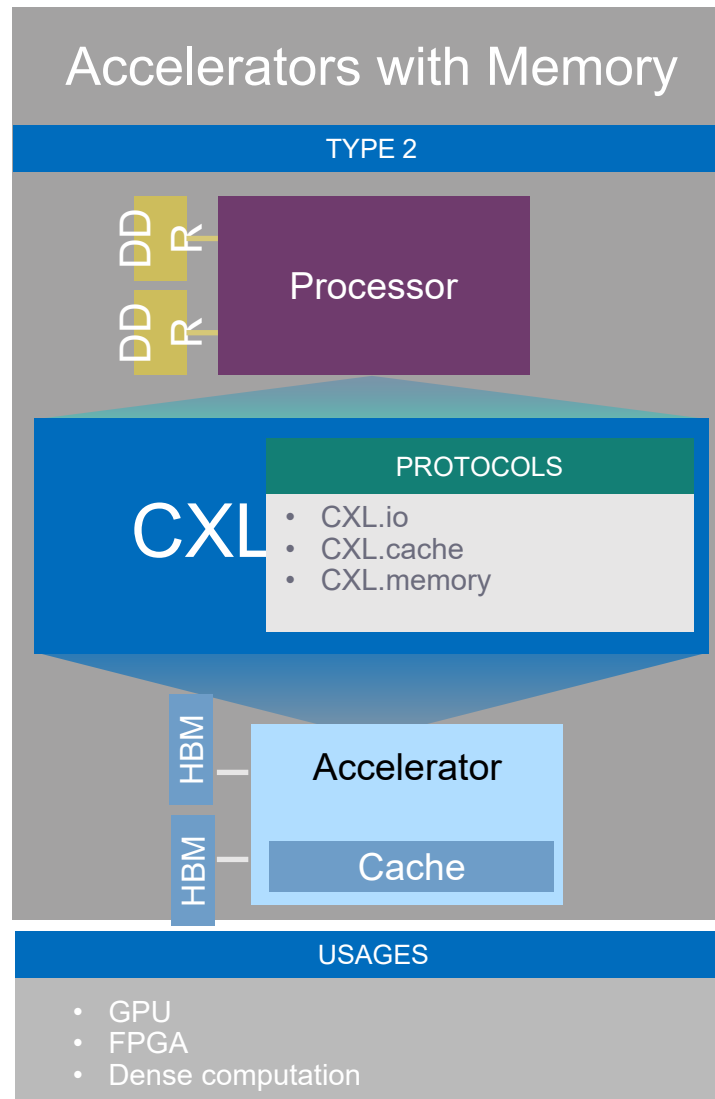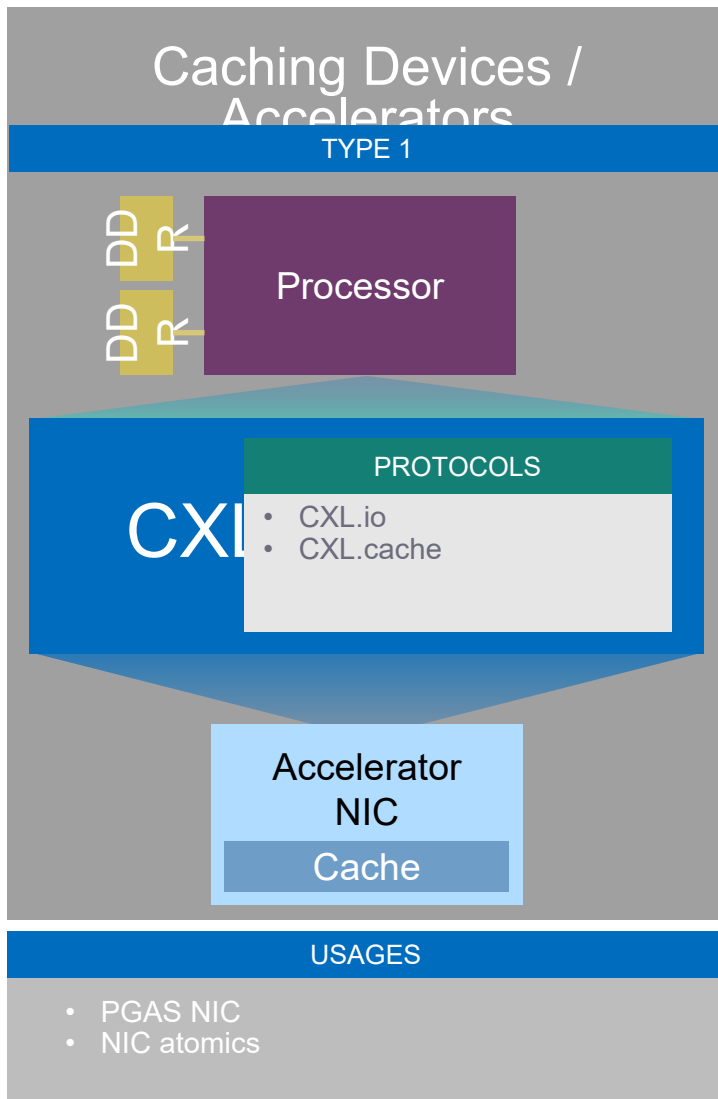| March 2019	September 2019	November 2020	August 2022
CXL 1.0 Specification Released	CXL 1.1 Specification Released	CXL 2.0 Specification Released	CXL 3.0 Specification Release

SNIA COMPUTE + MEMORY + STORAGE SUMMIT

Representative CXL Usages

Caching Devices / Accelerators	Accelerators with Memory	Memory Buffers
TYPE 1	**TYPE 2**	**TYPE 3**

Type 1

Processor

DDR / DDR

CXL

PROTOCOLS
- CXL.io
- CXL.cache

Accelerator NIC
Cache

USAGES
- PGAS NIC
- NIC atomics

Type 2

Processor

DDR / DDR

CXL

PROTOCOLS
- CXL.io
- CXL.cache
- CXL.memory

HBM

Accelerator
Cache

USAGES
- GPU
- FPGA
- Dense computation

Type 3

Processor

DDR / DDR

CXL

PROTOCOLS
- CXL.io
- CXL.memory

Memory / Memory / Memory / Memory

Memory Buffer

USAGES
- Memory BW expansion
- Memory capacity expansion
- **Persistent memory**

SNIA COMPUTE + MEMORY + STORAGE SUMMIT

CXL 3.0 Specification

Industry trends

- Use cases driving need for higher bandwidth include: high performance accelerators, system memory, SmartNIC and leading edge networking

- CPU efficiency is declining due to reduced memory capacity and bandwidth per core

- Efficient peer-to-peer resource sharing across multiple domains

- Memory bottlenecks due to CPU pin and thermal constraints

CXL 3.0 introduces…

- Fabric capabilities
  - Multi-headed and fabric attached devices
  - Enhance fabric management
  - Composable disaggregated infrastructure

- Improved capability for better scalability and resource utilization
  - Enhanced memory pooling
  - Multi-level switching
  - New enhanced coherency capabilities
  - Improved software capabilities

- Double the bandwidth

- Zero added latency over CXL 2.0

- Full backward compatibility with CXL 2.0, CXL 1.1, and CXL 1.0

SNIA COMPUTE + MEMORY + STORAGE SUMMIT

CXL 3.0 Specification

Fabric capabilities and management

Improved memory sharing and pooling

Enhanced coherency

Peer-to-peer

**Expanded capabilities for increasing scale and optimizing resource utilization**

- Fabric capabilities and fabric attached memory
- Enhance fabric management framework
- Memory pooling and sharing
- Peer-to-peer memory access
- Multi-level switching

- Near memory processing
- Multi-headed devices
- Multiple Type 1/Type 2 devices per root port
- Fully backward compatible to CXL 2.0, 1.1, and 1.0
- Supports PCIe® 6.0

SNIA COMPUTE + MEMORY + STORAGE SUMMIT

CXL 3.0 Spec Feature Summary

Features	CXL 1.0 / 1.1	CXL 2.0	CXL 3.0
Release date	2019	2020	August 2022
Max link rate	32GTs	32GTs	64GTs
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)			✓
Direct memory access for peer-to-peer			✓
Enhanced coherency (256 byte flit)			✓
Memory sharing (256 byte flit)			✓
Multiple Type 1/Type 2 devices per root port			✓
Fabric capabilities (256 byte flit)			✓

	Not supported
✓	Supported

SNIA COMPUTE + MEMORY + STORAGE SUMMIT

RECAP: CXL 2.0 Feature Summary
Memory Pooling
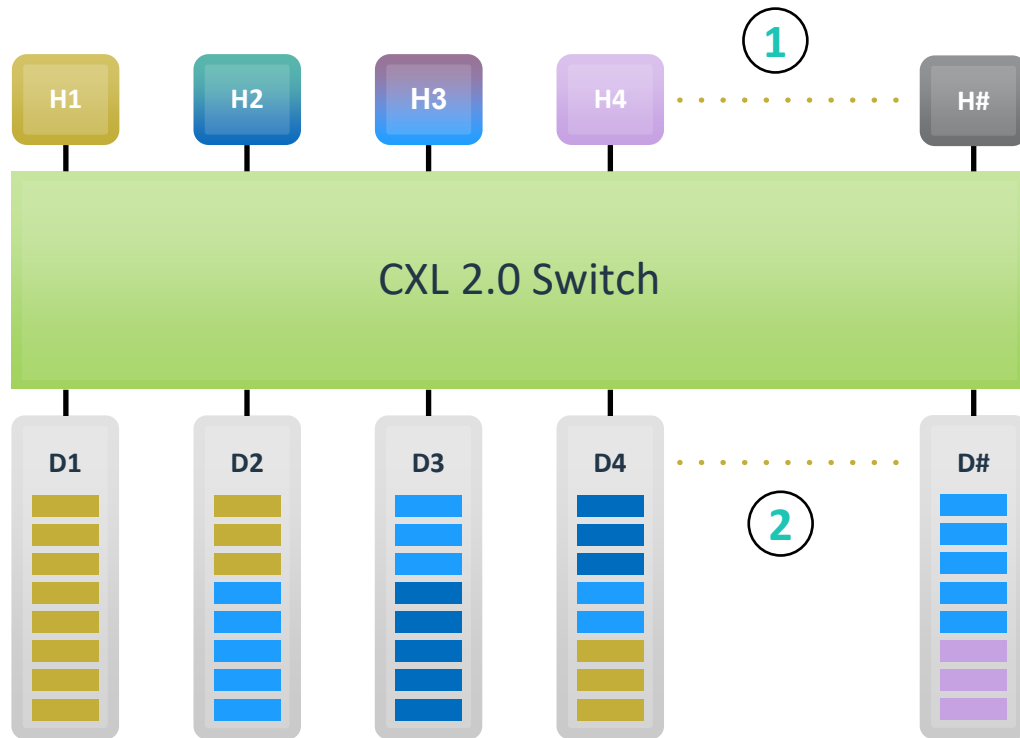
1. Device memory can be allocated across multiple hosts.

2. Multi Logical Devices allow for finer grain memory allocation
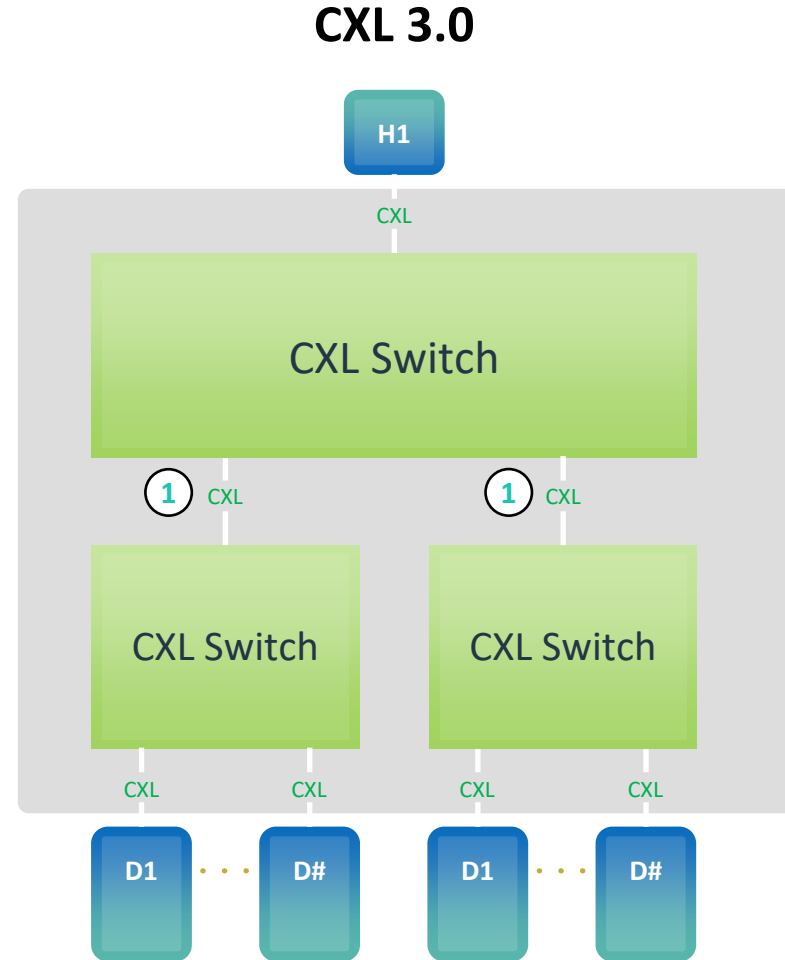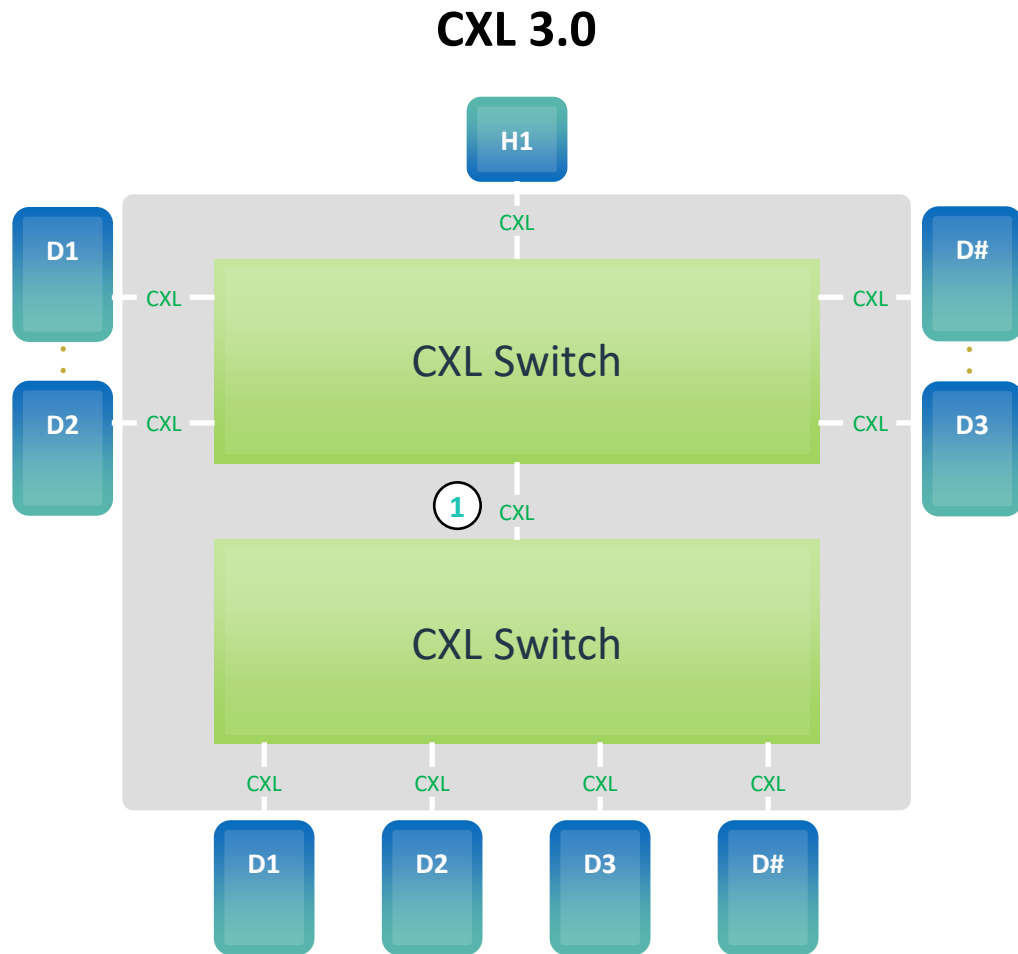
RECAP: CXL 2.0 Feature Summary
Switch Capability

CXL 2.0
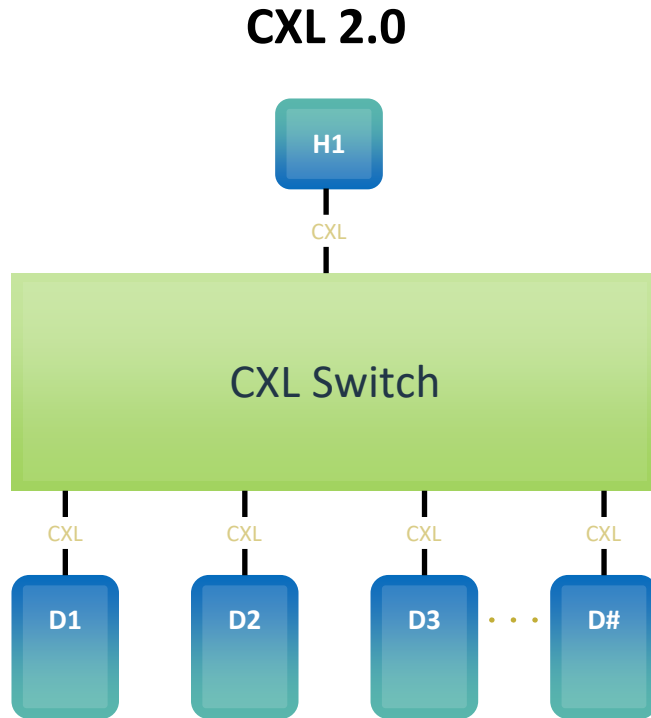
- Supports single-level switching

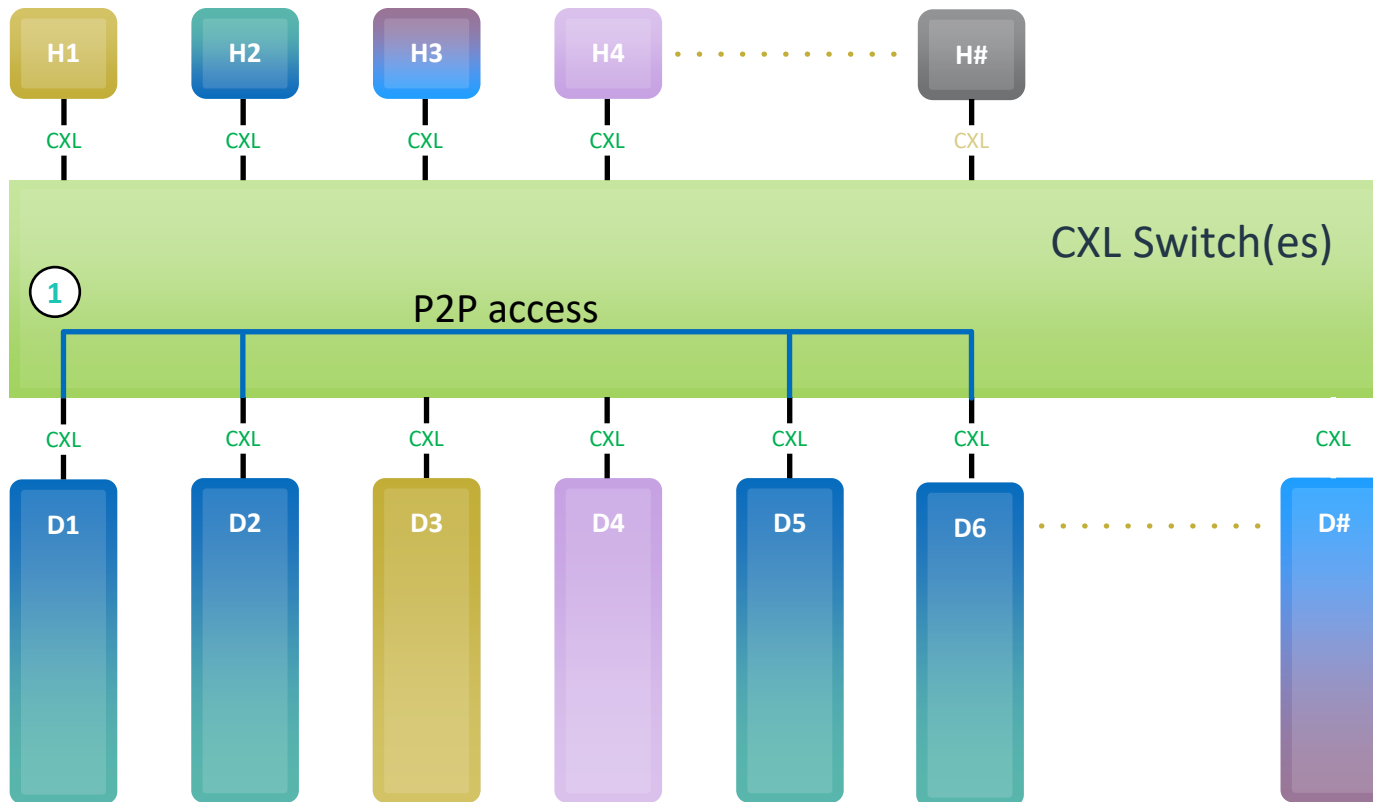- Enables memory expansion and resource allocation

CXL 3.0: Switch Cascade/Fanout
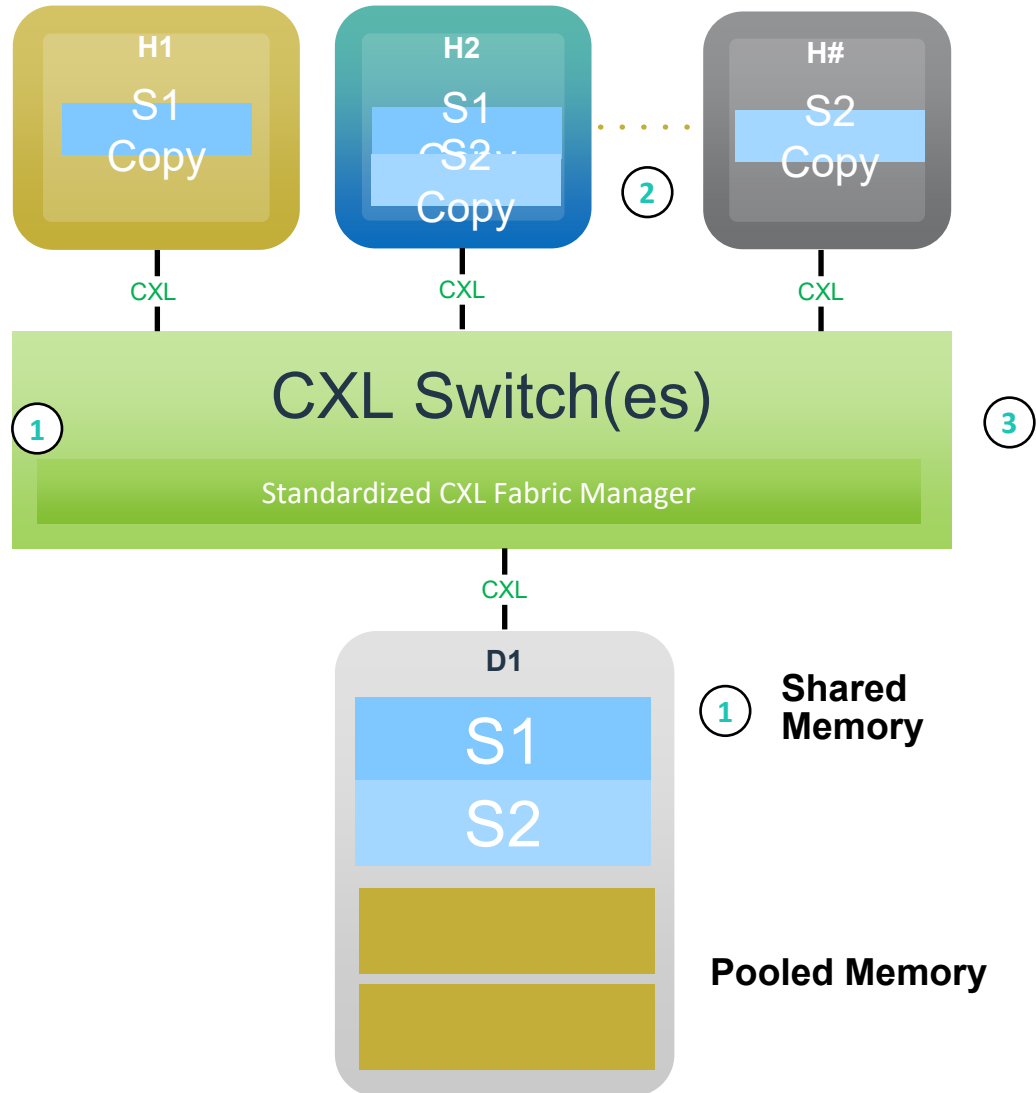## Supporting vast array of switch topologies

CXL 3.0: Device to Device Comms

① CXL 3.0 enables **peer-to-peer communication (P2P)** within a virtual hierarchy of devices

- Virtual hierarchies are associations of devices that maintains a coherency domain

CXL 3.0 Coherent Memory Sharing

1. Device memory can be shared by all hosts to increase data flow efficiency and improve memory utilization

2. Host can have a coherent copy of the shared region or portions of shared region in host cache

3. CXL 3.0 defined mechanisms to enforce hardware cache coherency between copies
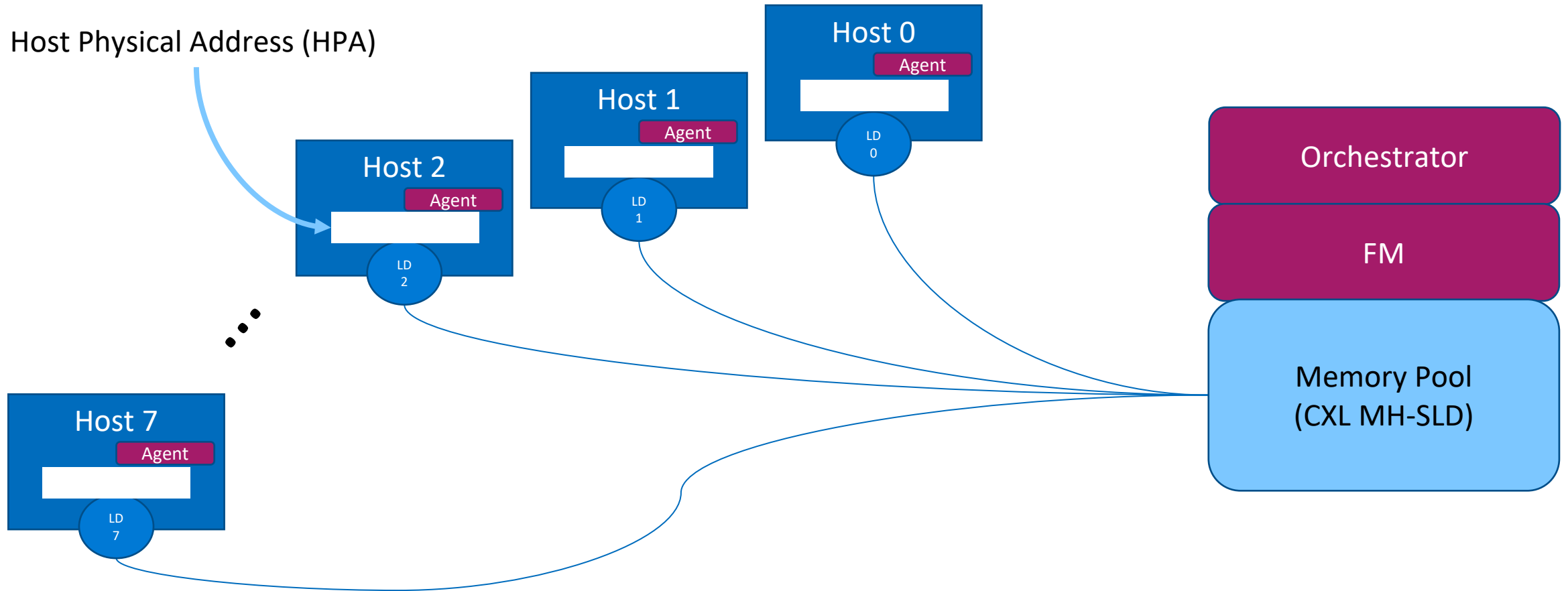
Dynamic Capacity Device (DCD)
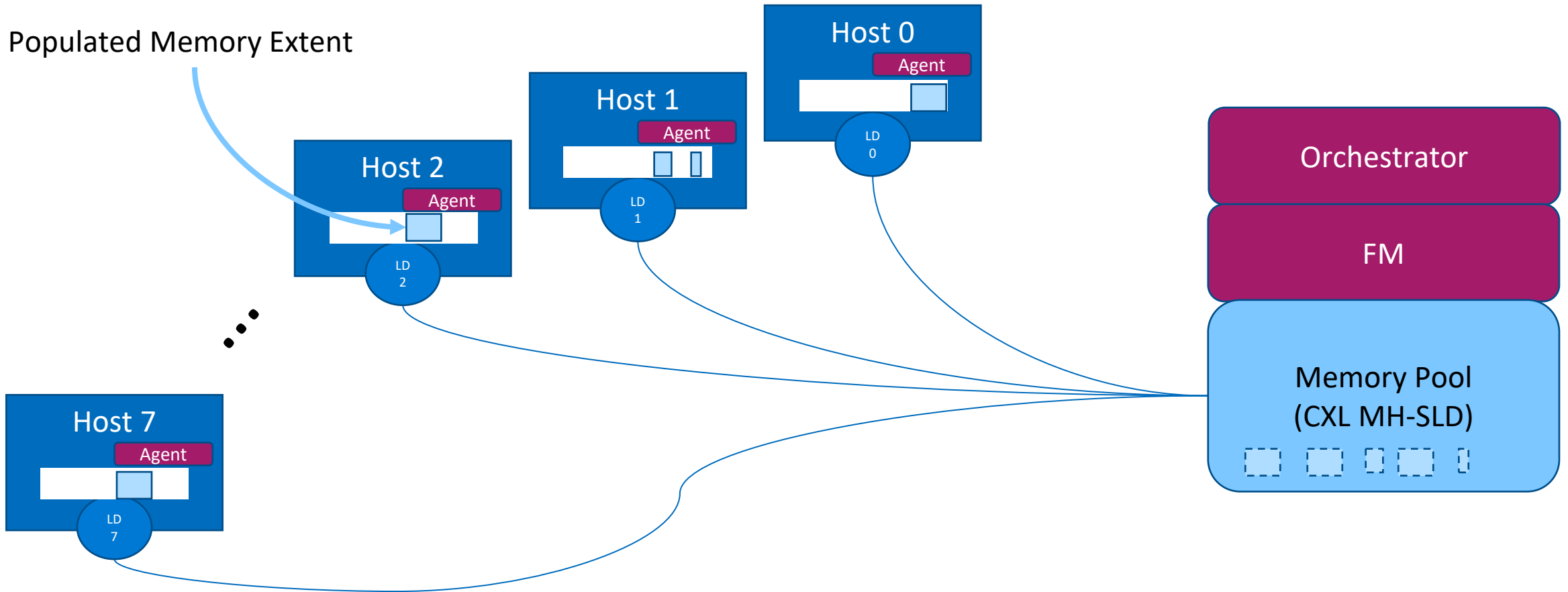
- Defined in CXL 3.0 Specification
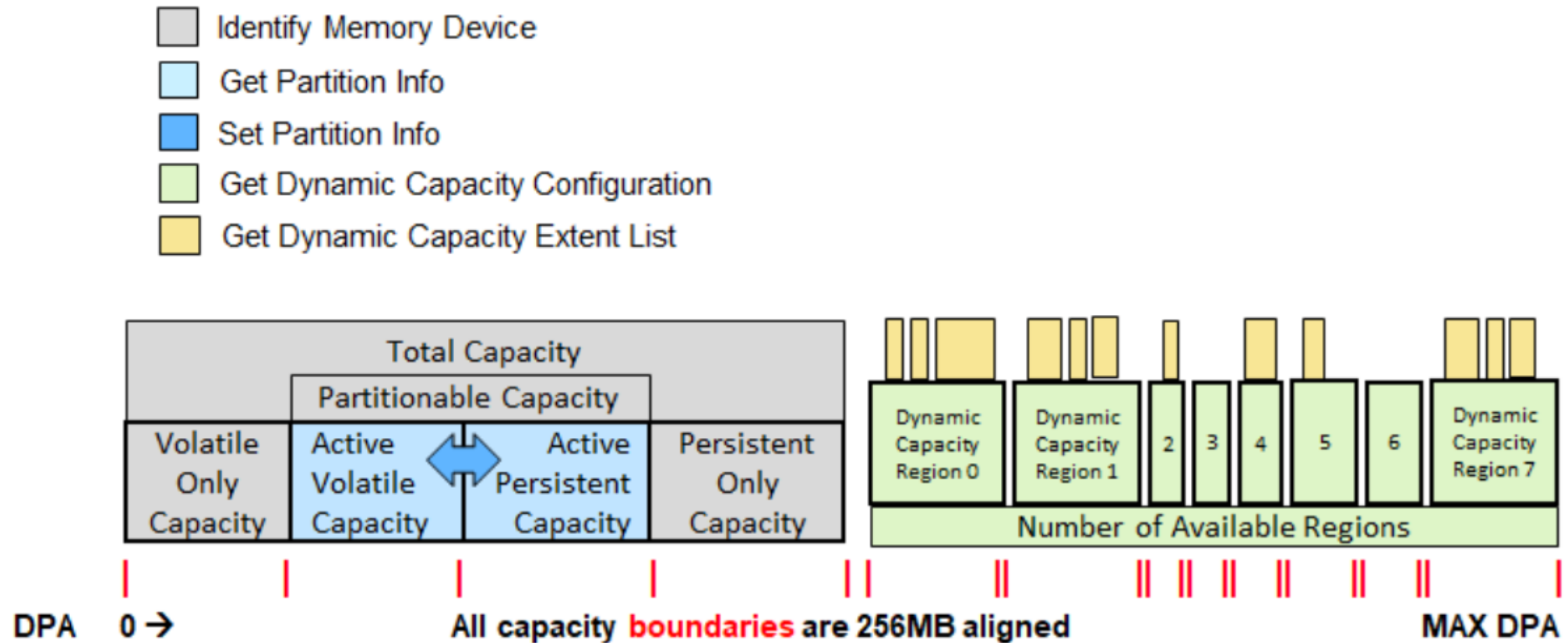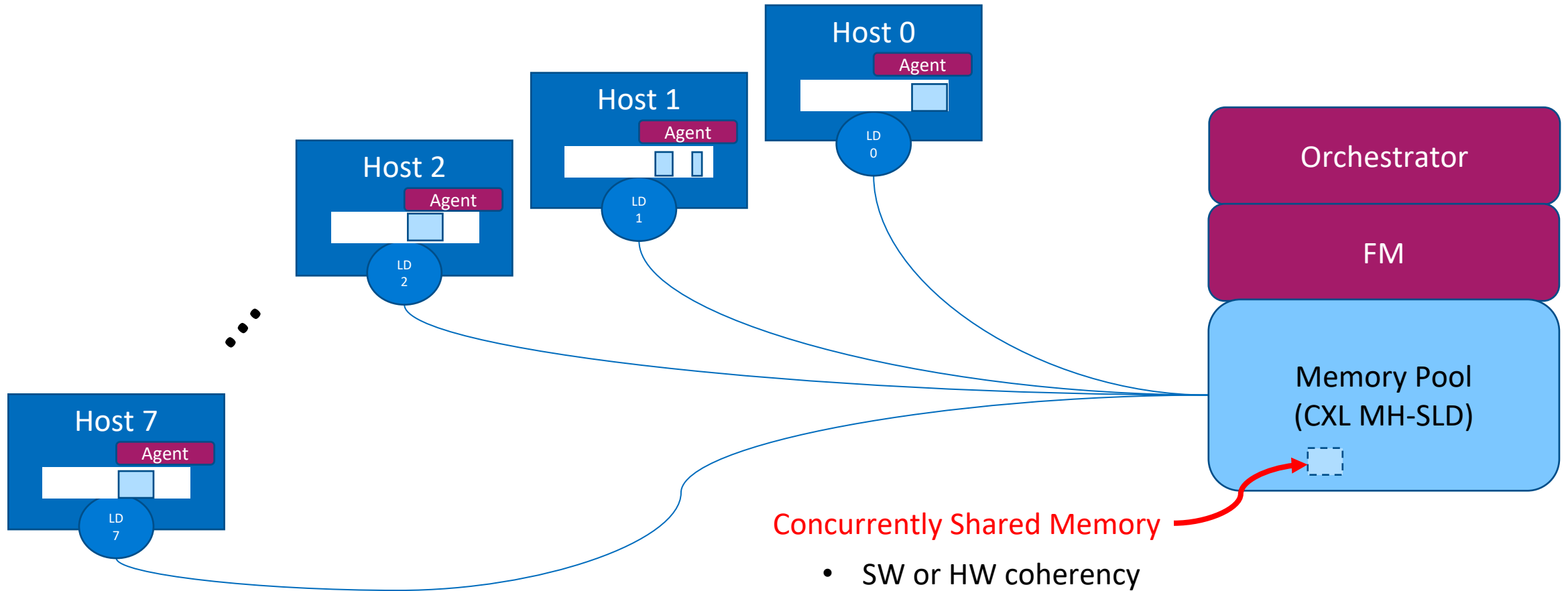
Example: Memory Pool
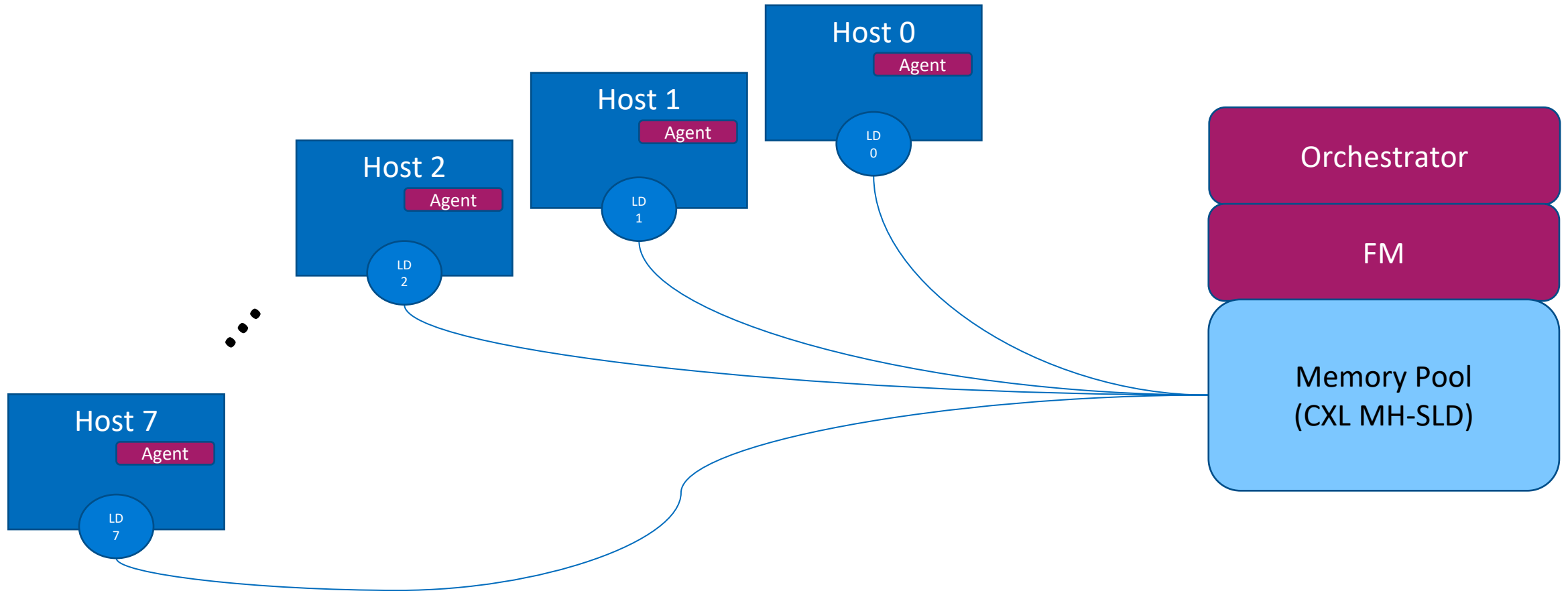
Example: Initial HDM Decoder Programming
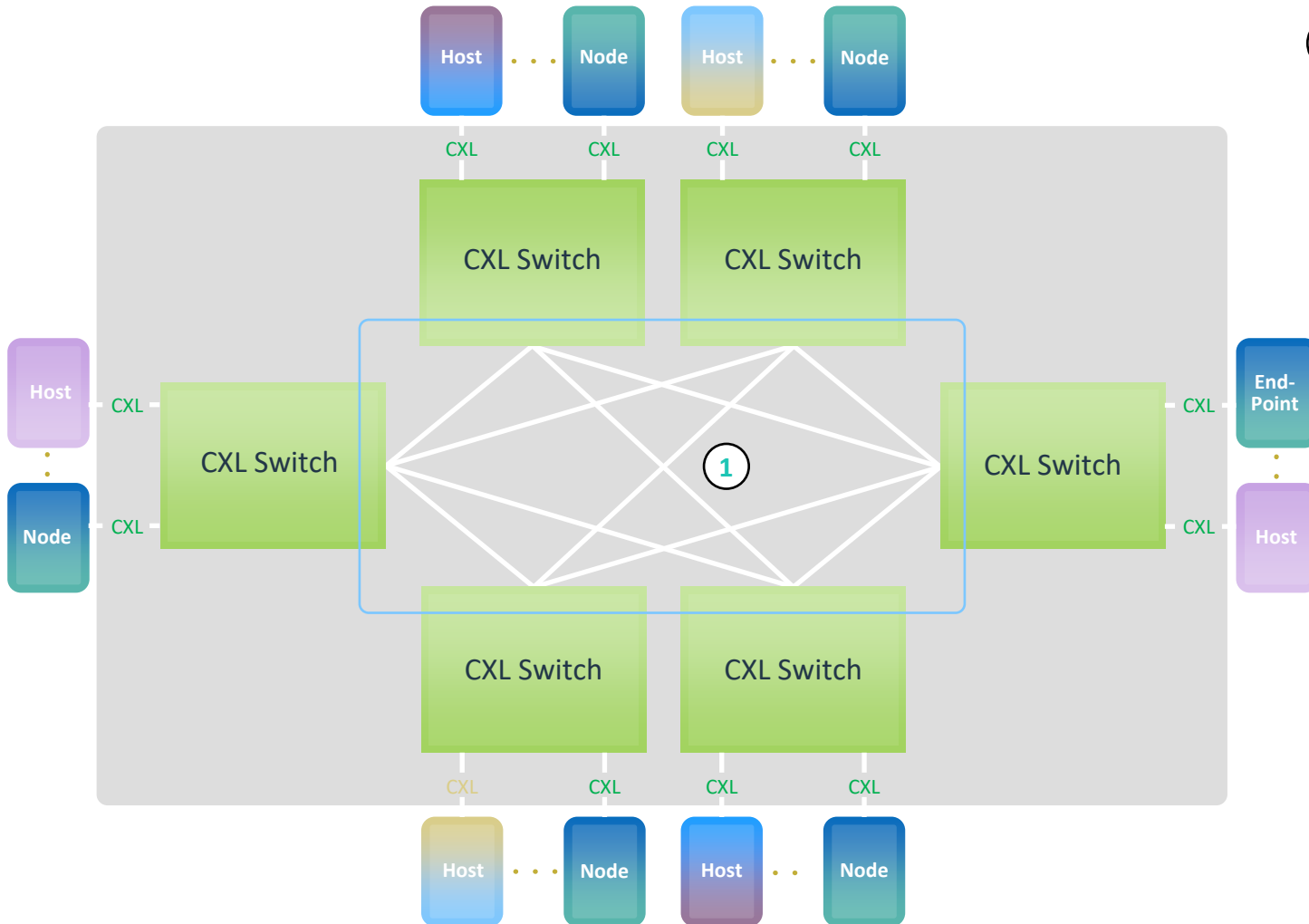
Example: Add Memory

Example: Shared Memory

Concurrently Shared Memory

- SW or HW coherency
- TBD: Application APIs for:
  - Allocation
  - Cross-host coordination

CXL 3.0: Fabrics Example

(1) Nodes can be **any combination**:

- Hosts
- Type 1 – Device with cache
  - Example: Smart NIC
- Type 2 – Device with cache and memory
  - Example: AI Accelerator
- Type 3 – Device with memory
  - Example: memory expander

See talk:
Introduction to Fabrics in CXL 3.0
by Vince Haché

CXL 3.0 Summary

- **CXL 3.0 features**
  - Full fabric capabilities and fabric management
  - Expanded switching topologies
  - Enhanced coherency capabilities
  - Peer-to-peer resource sharing
  - Double the bandwidth and zero added latency compared to CXL 2.0
  - Full backward compatibility with CXL 2.0, CXL 1.1, and CXL 1.0

- **Enabling new usage models**
  - Memory sharing between hosts and peer devices
  - Support for multi-headed devices
  - Expanded support for Type-1 and Type-2 devices
  - GFAM provides expansion capabilities for current and future memory

- **Call to Action**
  - Download the CXL 3.0 specification
  - Support future specification development by joining the CXL Consortium
  - Follow us on Twitter and LinkedIn for updates!

SNIA COMPUTE + MEMORY + STORAGE SUMMIT

Please take a moment to rate this session.

Your feedback is important to us.