Compute Express Link™ (CXL™): Advancing the Next Generation of Data Centers

Alan Benjamin, CEO, GigaIO

Industry Landscape

Proliferation of
Cloud Computing

Growth of
AI & Analytics

Cloudification of
the Network & Edge

Alibaba Group
AMD
arm
cisco
DELL EMC
Google
Hewlett Packard Enterprise
HUAWEI
IBM
intel
Meta
Micron
Microsoft
NVIDIA
SAMSUNG

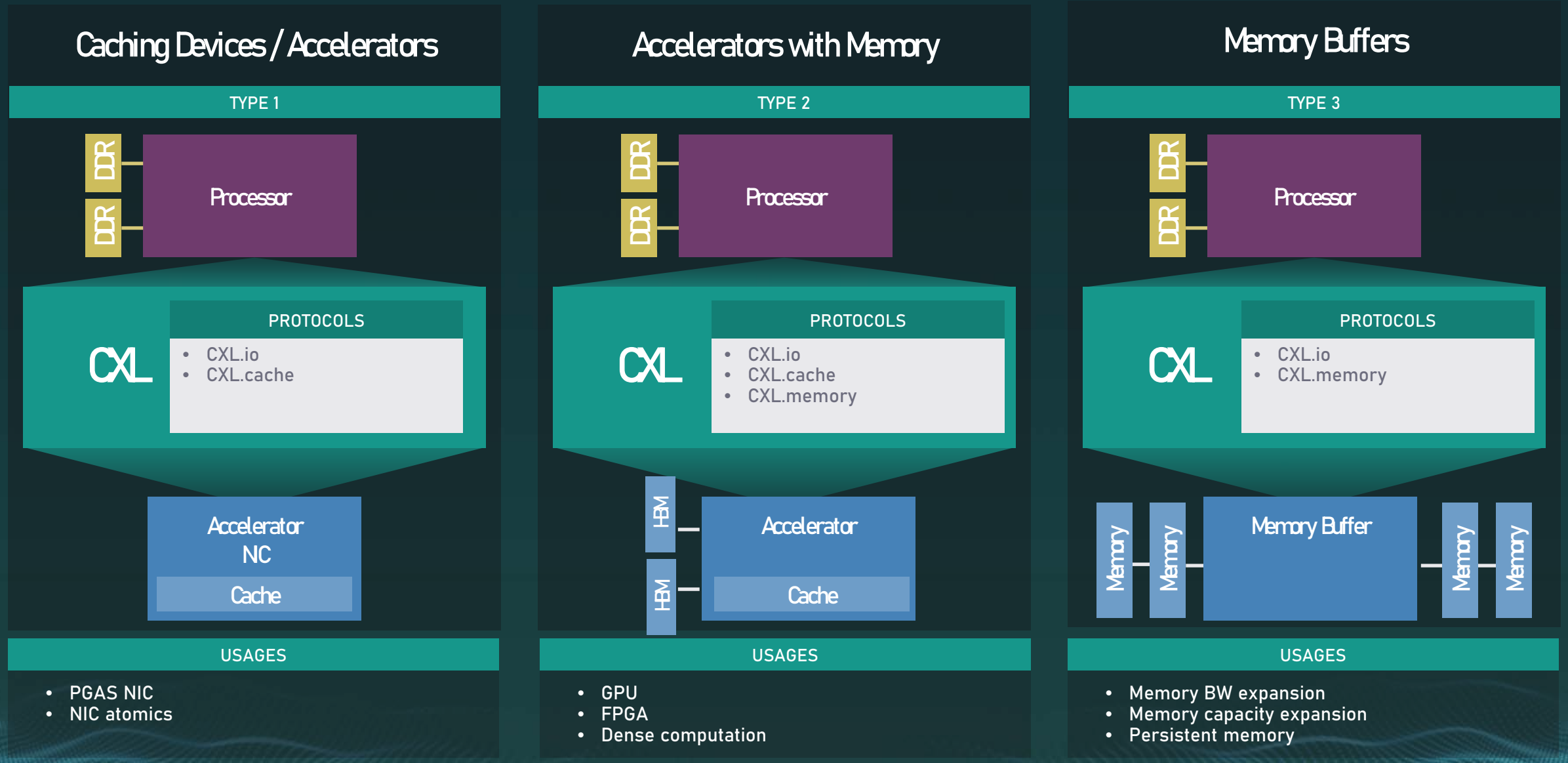
CXL Board of Directors

CXL Compute Express Link™

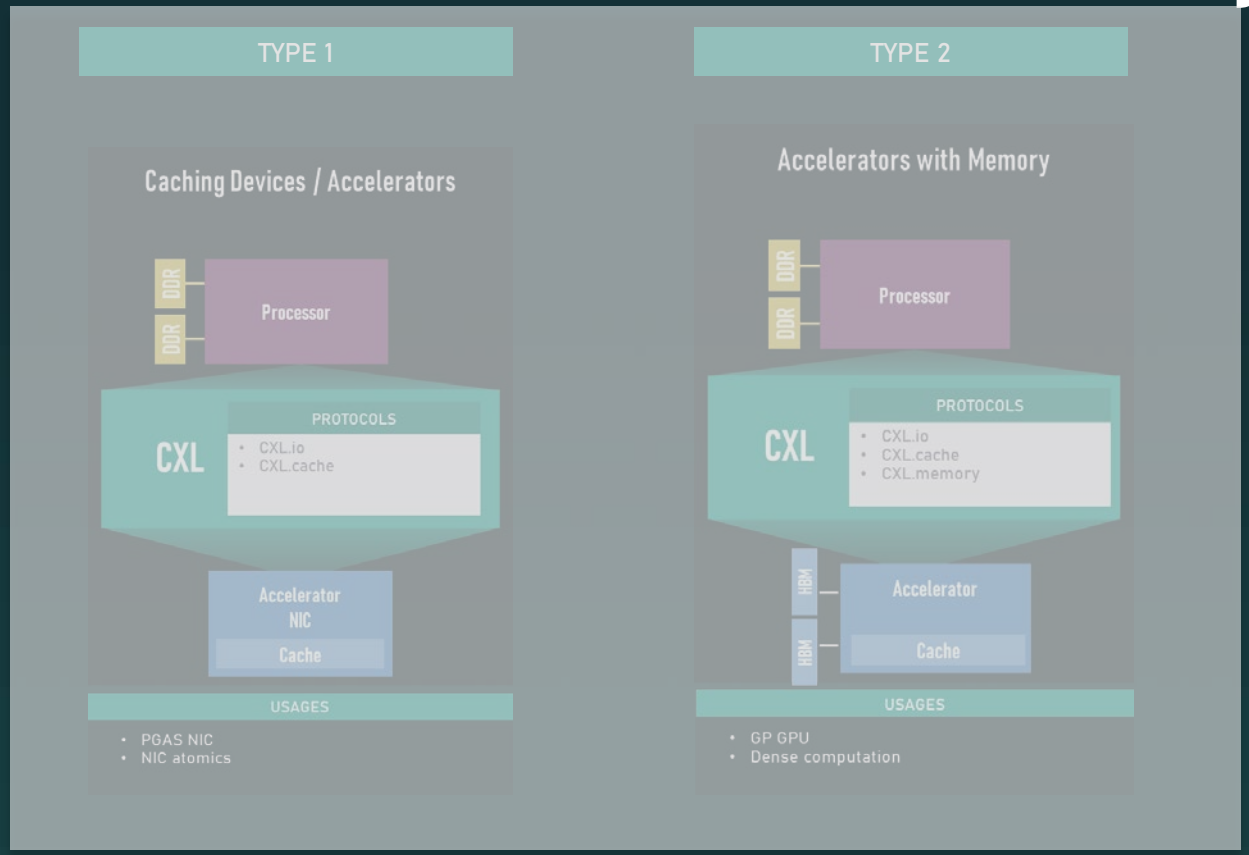
Industry Open Standard for High Speed Communications

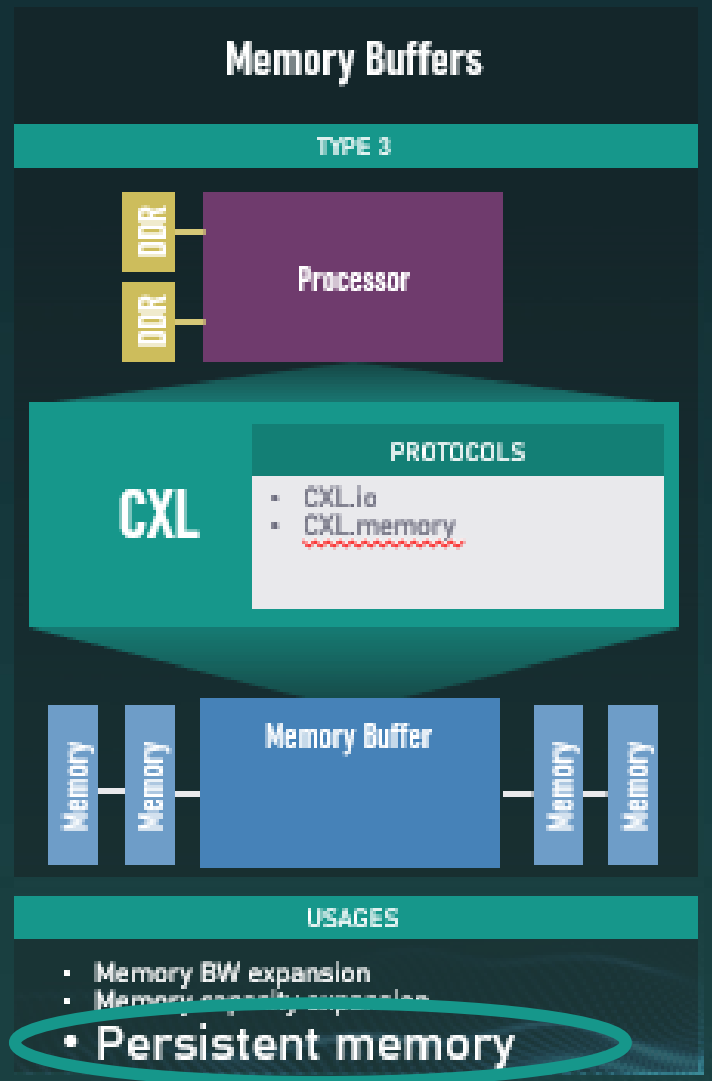
180+ Member Companies

CXL 2.0 Usage Models – Recap



Persistent Memory is a Key CXL Usage



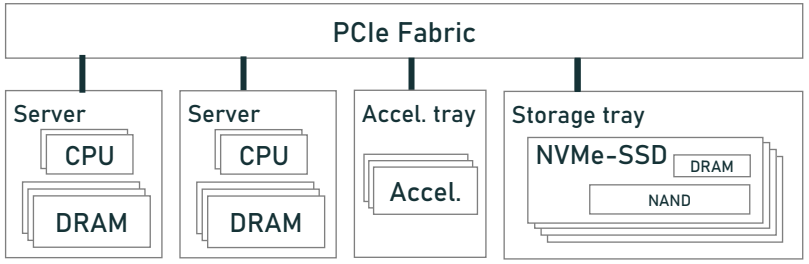


What is the Fabric Technology for Disaggregated & Heterogeneous Computing System?

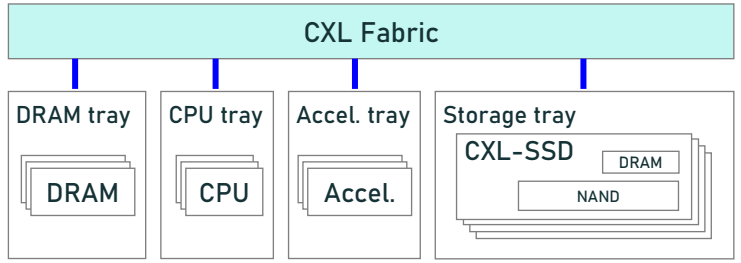
- CXL fabric can be a key element for composable DC infrastructure
  - CXL.io for flexible path / CXL.mem for short latency & high bandwidth / CXL.cache for tightly aligned to the host
- Peer-to-Peer function helps system to enjoy higher bandwidth and lower latency with CXL based devices

Transition of Fabric Technology: PCIe → CXL

PCIe-based Rack-Scale Computing
scaling beyond enclosure, thanks to:
memory semantic and P2P



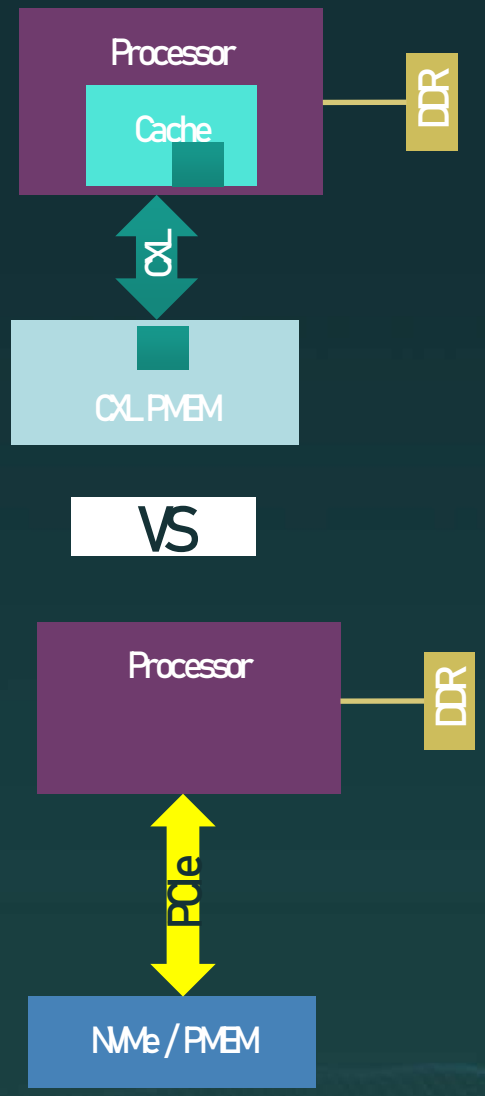
CXL-based Rack-Scale Computer
easier scaling beyond enclosure, thanks to:
cache coherency and memory management



Persistent Memory (PMEM)

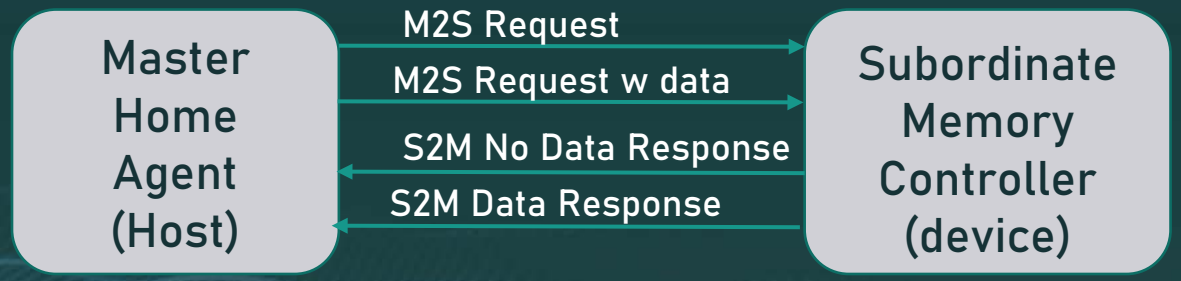
- Characteristics and benefits of Persistent Memory (PMEM)
  - Byte-addressable (vs. NVMe is block addressable)
  - Generally, lower latencies compared to SSD
  - Cacheable (vs. NVMe is uncached)
  - Data persists across power loss (vs. DRAM loses its content)
  - Generally, larger capacity (vs. DRAM)

- Many Workloads benefit from PMEM
  - Traditional Databases – Accelerated logging/journaling, instant recovery
  - Analytics/AI/ML – real time access to large datasets, faster checkpointing
  - Storage – caching, tiering, ..
  - HPC – Reduce checkpointing overhead
  - and more ..



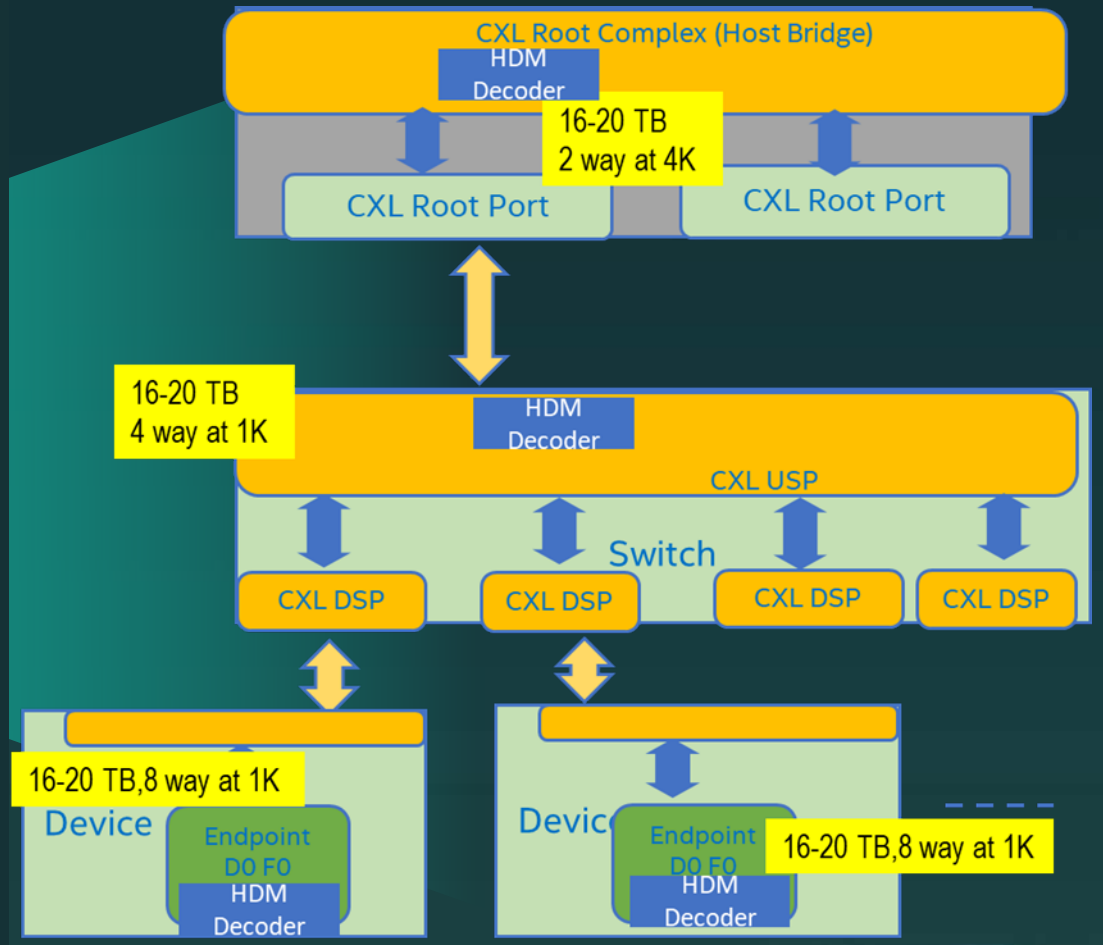
CXL Protocol is Well Suited for PMEM

- CXL.mem protocol is transactional, see below
  - PMEM media may have longer latencies, or variable access latencies
  - Controller can hide longer and/or variable access latencies
- CXL.mem abstraction
  - Memory Controller and media are abstracted
  - Enables new and innovative media types
- CXL 2.0 introduces memory QoS
  - Device can synchronously report how loaded it is
  - Can prevent head of line blocking in heterogenous memory configuration (e.g. DRAM + PMEM)
- CXL 2.0 adds Memory Interleaving, a standardized register interface and Global Persistent Flush (GPF)



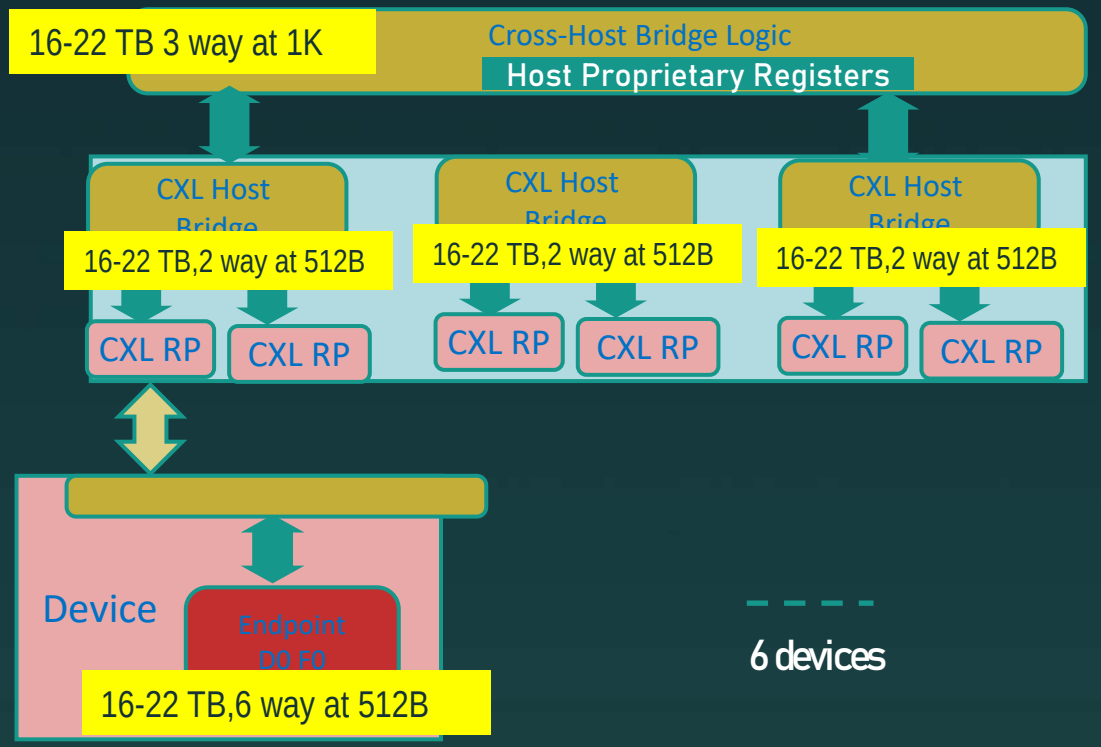
CXL Supports Interleaving of Memory Devices

- Interleaving is performance feature
- Example: 8-way interleaved device
- How it works
  - CXL Host Bridge HDM Decoders configured to select one of two Root Ports based on A[12]
  - HDM Decoders in every switch configured to select one of four DSPs based on A[11:10]
  - HDM Decoders in the device configured for 8 way interleave at 1K
  - Device removes A[12:10] from the Host Physical Address when computing the Device Physical Address



Memory Interleaving Enhancements

- CXL 2.0 specification defined 2, 4 and 8-way interleaving options
- This ECN introduces 3, 6, 12 and 16-way interleaving options
  - Matches interleaving options that are available with native DDR
  - Sweet spots that CXL 2.0 spec missed
- No impact to Switches
- 3, 6 and 12-way interleaving implies 3-way math
  - Hosts support mod3 to select one out of 3 (or 6 or 12) targets
  - The device needs to implement divide by 3 (or 6 or 12) operation on HPA
- The ECN limits the legal combinations
  - Limits scope
  - Avoid potential address aliasing when mixing 2 way and 3-way math
- 6 way interleaving across 6 Root Ports shown here



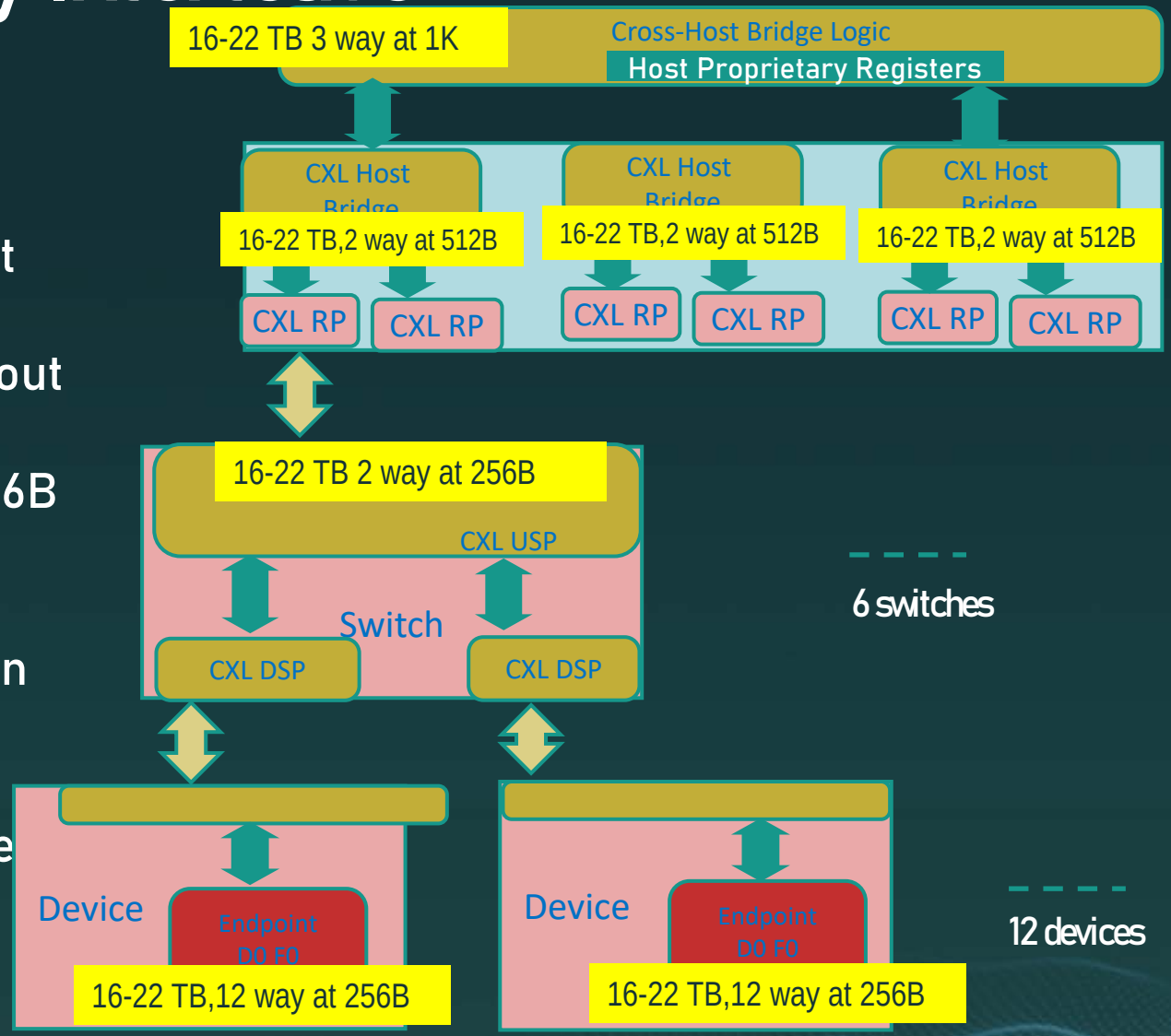
12-way Interleave

- A more complex, 12-way interleaving example shown here

- Cross-host Bridge logic picks one CXL Host Bridge out of 3 at 1K granularity
- Each CXL Host Bridge picks one Root Port out of 2 at 512B Granularity
- Each USP selects one DSP our of two at 256B granularity

- The device performs a divide by 12 operation during HPA to DPA conversion

- Contiguous address bits used for interleave selection



A sneak peak into the CXL 3.0 Specification

Industry trends

- Use cases driving need for higher bandwidth include: high performance accelerators, system memory, SmartNIC and leading edge networking
- Interconnect needed that can optimize system level flows among those components with advanced switching, efficient P2P and fine-grained resource sharing across multiple domains

CXL 3.0 features will have…

- Double bandwidth
- Improved capability for better scalability and improved resource utilization
  - Enhanced memory pooling and enables new memory usage models
  - Multi-level switching with multiple host and fabric capabilities and enhanced fabric management
  - New symmetric coherency capabilities
  - Improved software capabilities
- Fully backward compatible with CXL 1.0, CXL 1.1, CXL 2.0
- Specification available 2022

In Summary

CXL Consortium momentum continues to grow

- 180+ members and growing
- Responding to industry needs and challenges

CXL is ideal for attaching Persistent Memory

- The protocol designed with PMEM in mind, media-agnostic
- Generic driver model eases SW enabling
- Robust RAS and reliability features
- Variety of form factors enable innovative system designs

Call to action

- Join CXL Consortium
- Follow us on YouTube, Twitter and LinkedIn for more updates!

CXL Compute Express Link™

Thank You