

# Introducing the CXL 4.0 Specification

December 4, 2025



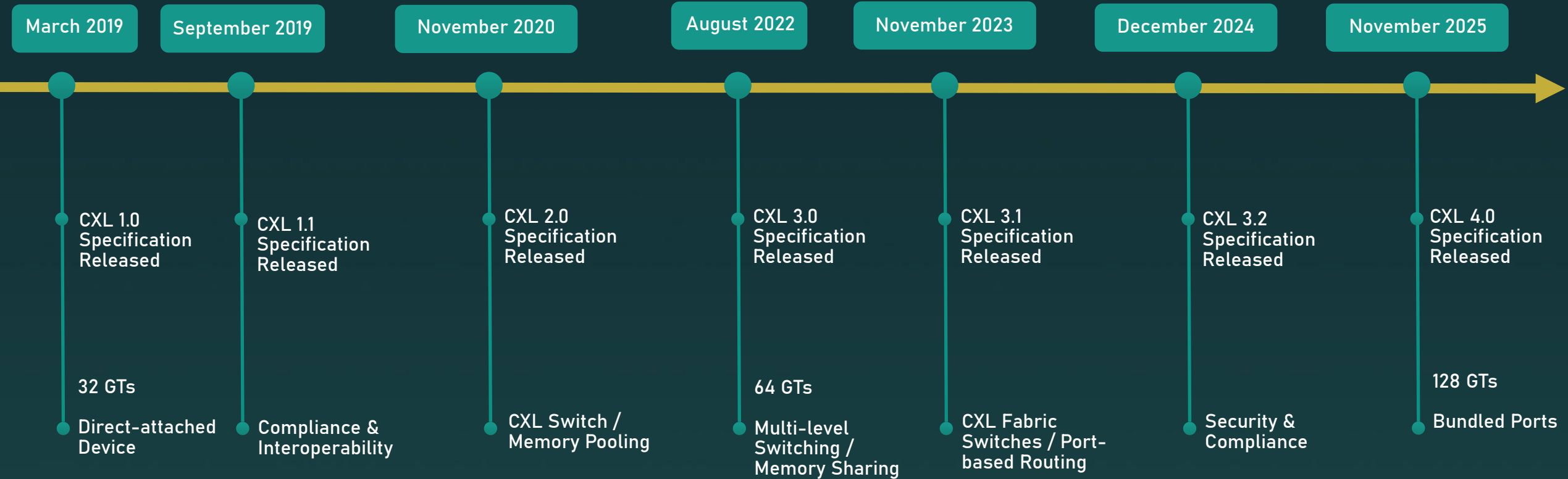
CXL Board of Directors



Industry Open Standard for  
High Speed Communications

280+ Member Companies

# CXL Specification Release Timeline



- Industry Landscape and CXL
- Computing challenges addressed by CXL (3 generations)
- CXL 4.0 Features
- Conclusions



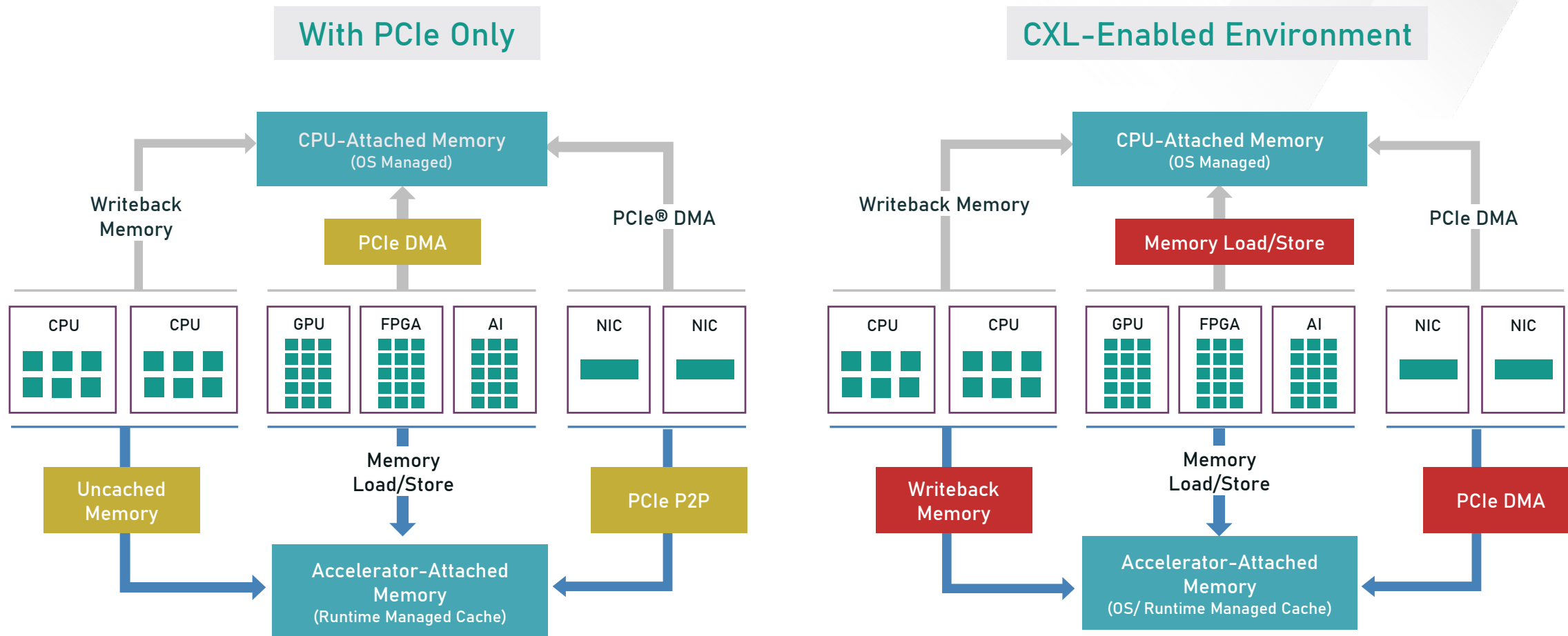
# Industry Trends

- AI and ML applications on cloud and on-premises
- Memory-intensive applications (database, fraud-detection, animation, retail etc.)
- Compute and silicon diversification (Heterogeneous computing)
- Disaggregation of memory from compute (memory pooling/sharing, etc.)
- Media Agnostic memory tiers deployed to decrease overall platform costs



- Industry Landscape and CXL
- **Computing challenges addressed by CXL (3 generations)**
- CXL 4.0 Features
- Conclusions

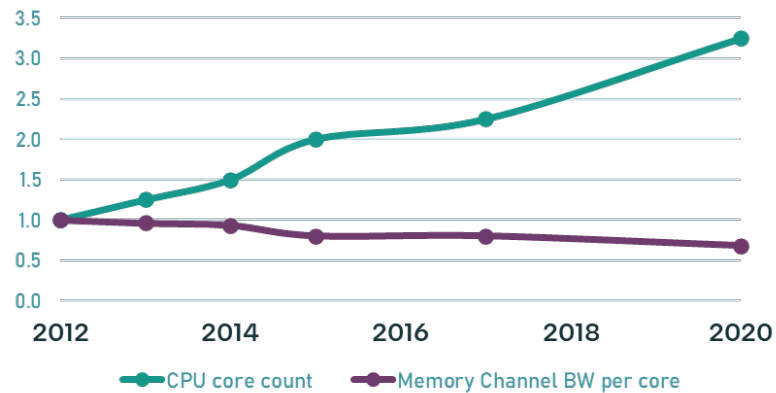
# CXL®: Heterogeneous Compute – Challenge 1



Coherency and Memory Semantics added on PCIe Infrastructure

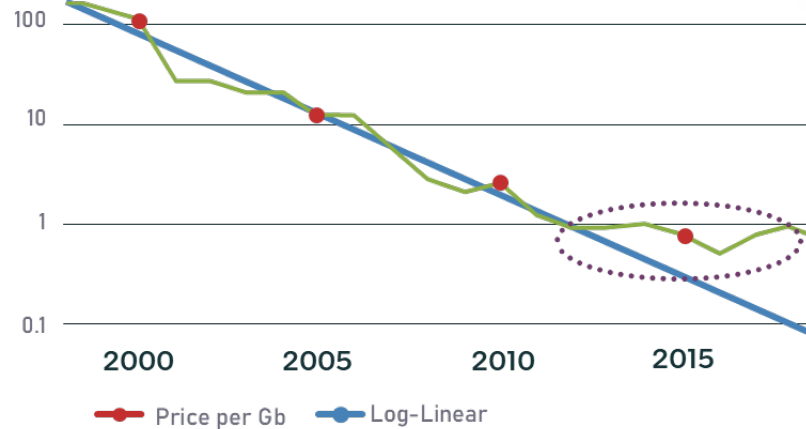
# System Memory Scalability – Challenge 2

Normalized Growth Rate



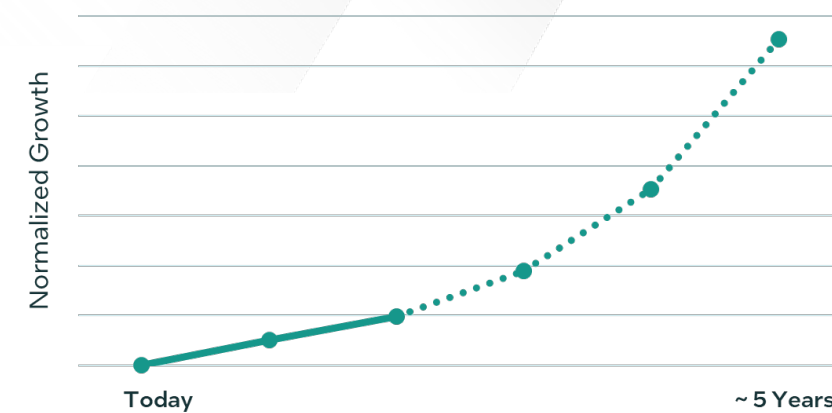
- Increasing core counts drives memory demand
- Increasing bandwidth and capacity
- Memory is not able to keep up -> more DDR channels (cost, power and feasibility challenges)

Price per Gb (Log Scale)



- Memory is an increasing % of system power and cost
- Memory price (cost/bit) is flat due to scaling challenges
- Memory power scaling with speed

Compute Demand Exploding



AI/ ML Models are Growing Rapidly:

- ~50x growth in ~5 years
- Existing memory hierarchy can't keep pace

DDR5-6400 offers 50 GB/s with ~200 signal-pins. A x16 PCIe 6.0 at 64.0 GT/s offers 512 GB/s raw with 64 signal pins with no restriction of 15W per DIMM!

Challenge 2 addressed by CXL: Memory capacity and bandwidth expansion with scalability



# CXL Approach

## Coherent Interface

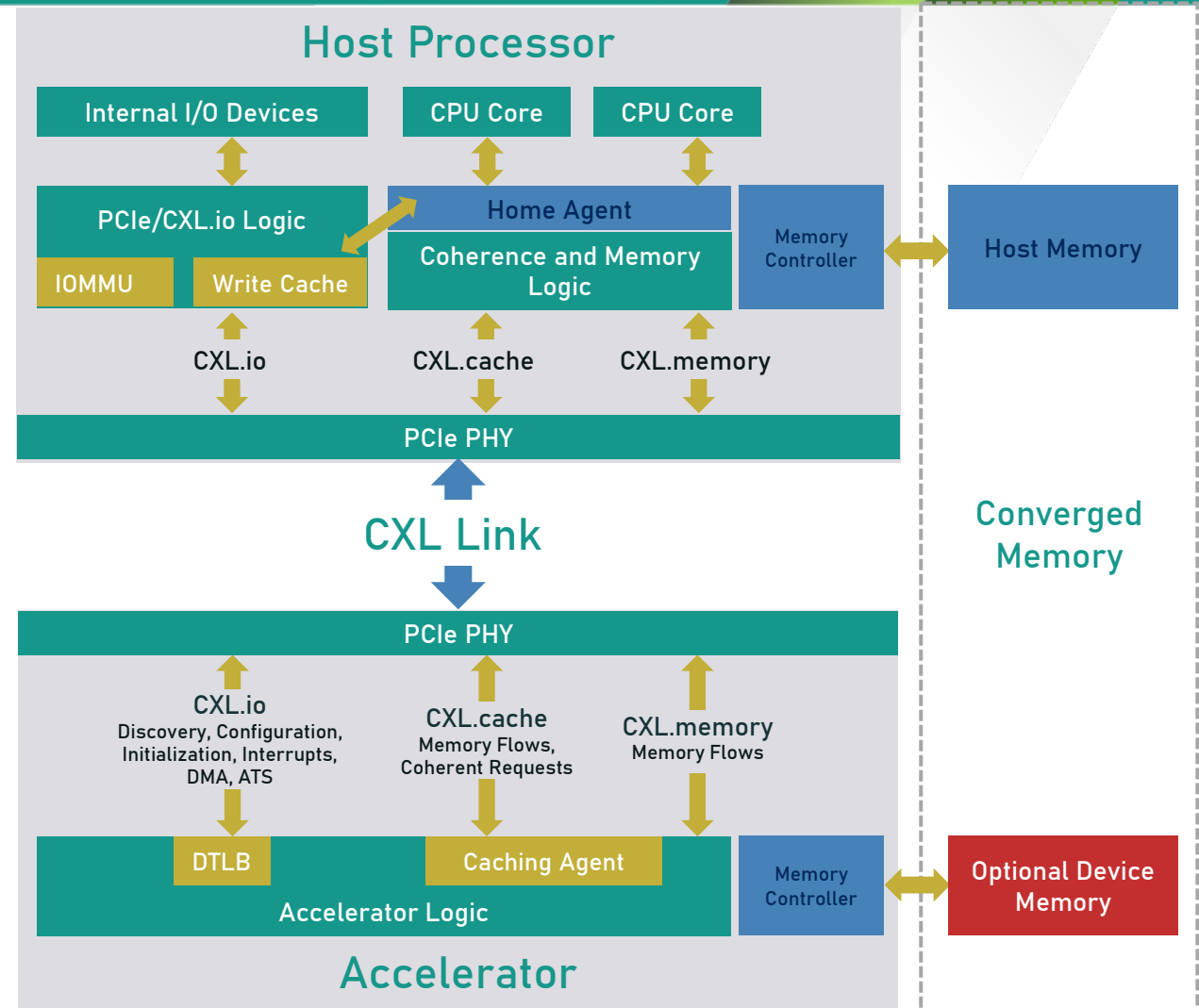
- Leverages PCIe with three multiplexed protocols
- Built on top of PCIe® infrastructure

## Low Latency

- CXL Cache/CXL Memory targets near CPU cache coherent latency (<200ns load to use)

## Asymmetric Complexity

- Eases burdens of cache coherence interface designs for devices



Building on this approach for backwards-compatible evolution – to 4<sup>th</sup> generation and beyond while including new usage models

# CXL 1.0/CXL 1.1 Usage Models w/ Direct Connect

## Type 1 Device

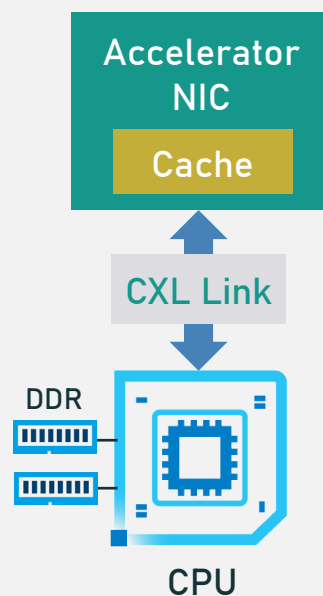
Caching Devices/Accelerators

Usages:

- PGAS NIC
- NIC atomics

Protocols:

- CXL.io
- CXL.cache



## Type 2 Device

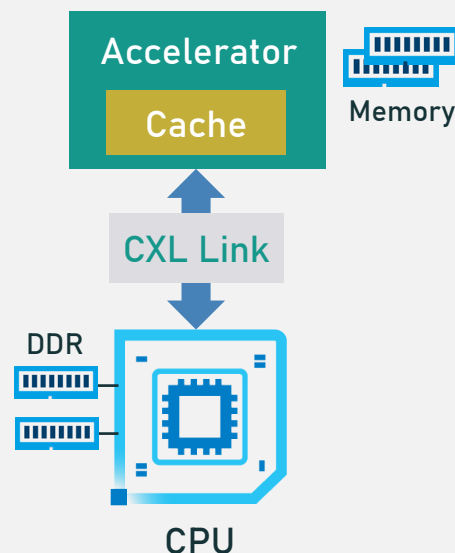
Accelerators with Memory

Usages:

- GPU
- FPGA
- Dense Computation

Protocols:

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



## Type 3 Device

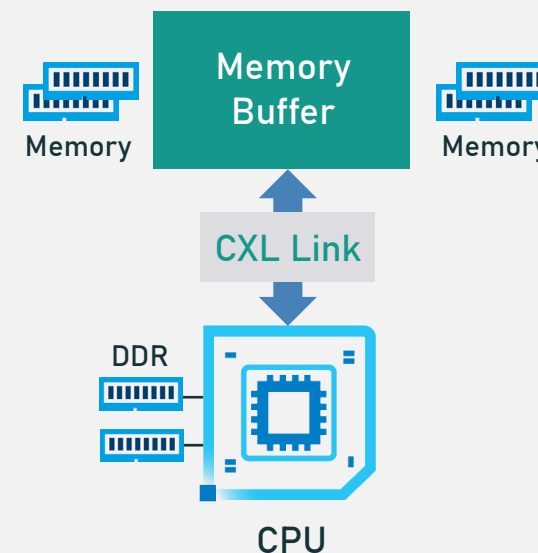
Memory Buffers

Usages:

- Memory BW expansion
- Memory capacity expansion
- 2LM

Protocols:

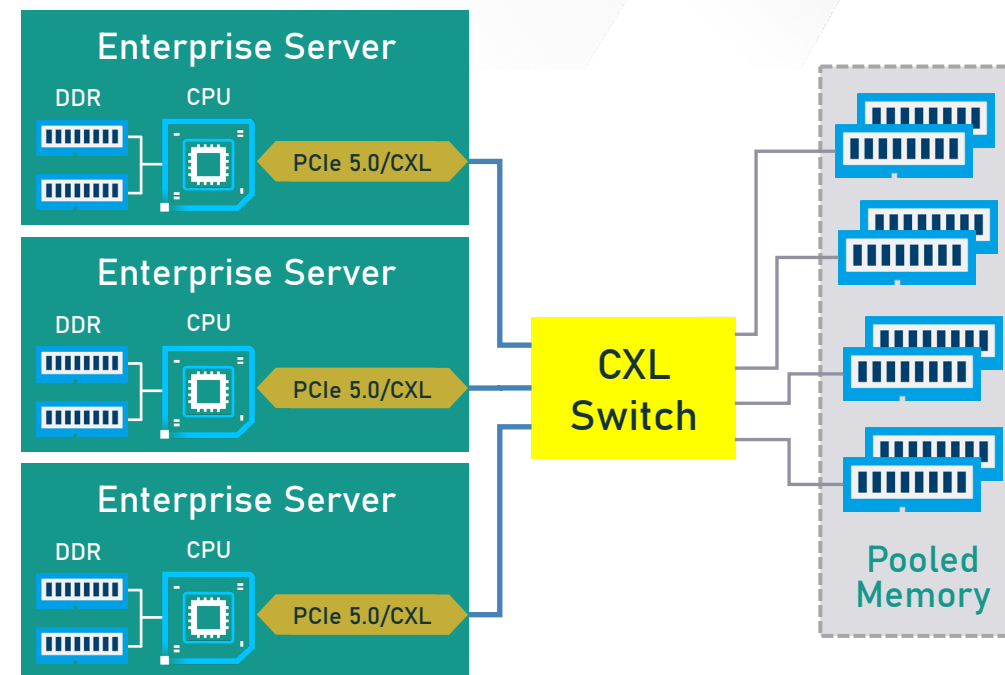
- CXL.io
- CXL.mem



Challenges addressed: (1) Heterogeneous compute and (2) Memory scalability

# CXL 2.0: Resource Pooling at Rack Level, Persistent Memory Support and Enhanced Security

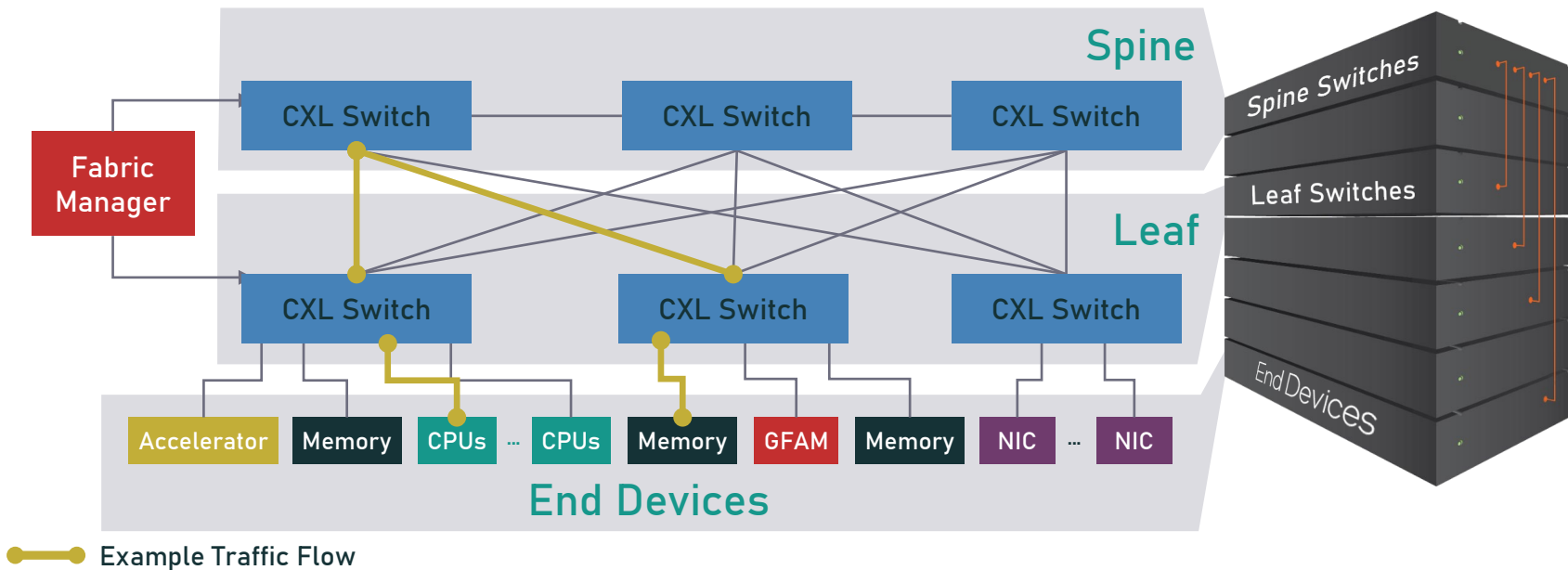
- Resource pooling/disaggregation
  - Managed hot-plug flows to move resources
  - Type-1/Type-2 device assigned to one host
  - Type-3 device (memory) pooling at rack level
  - Direct load-store, low-latency access – similar to memory attached in a neighboring CPU socket (vs. RDMA over network)
- Hot-plug; On/Off-lining support
- Persistence flows for persistent memory
- Fabric Manager/API for managing resources
- Security: authentication, encryption
- Beyond node to rack-level connectivity!



Challenge 3: Stranded memory and compute resources at Data Centers.  
Disaggregated system with CXL optimizes resource utilization delivering lower TCO and power efficiency

# CXL 3.0 Enhancements

- Bandwidth doubling with 64 GT/s at 0-latency add
- Protocol enhancements with direct peer-to-peer to HDM memory
- Composable systems with spine/leaf architecture at rack/pod, Scale-out Fabric (PBR)
- Shared Memory
- Confidential Compute



## CXL 3.0 Fabric Architecture

- Interconnected spine switch system
- Leaf switch NIC enclosure
- Leaf switch CPU enclosure
- Leaf switch accelerator enclosure
- Leaf switch memory enclosure

Challenge 4: Fine-grained data sharing/ message passing in distributed composable systems

# CXL 4.0 Features

# Key Feature Enhancements for CXL 4.0



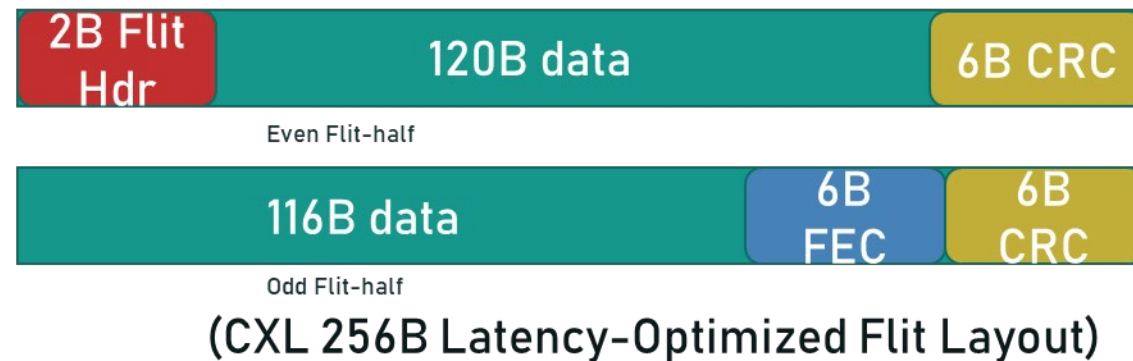
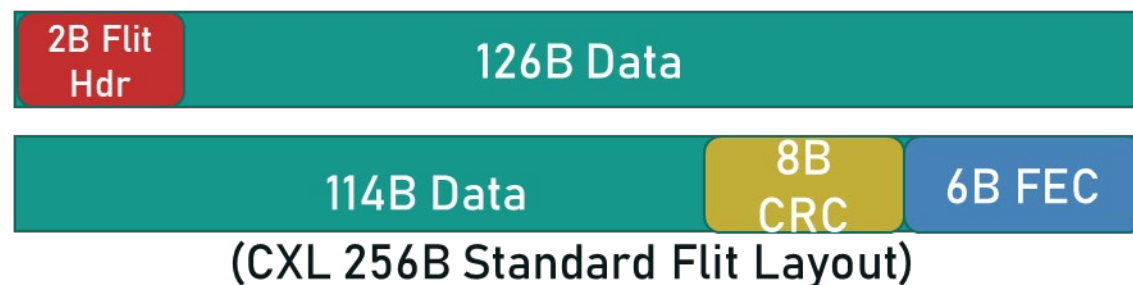
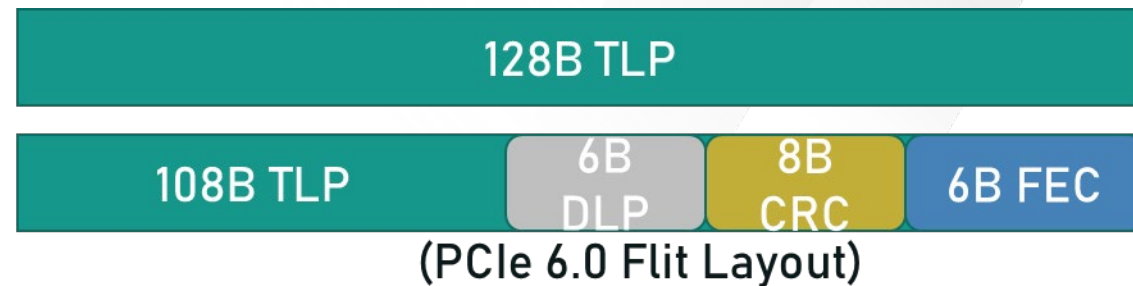
- Doubles the bandwidth to 128GTs with zero added latency
  - Enables rapid data movement between CXL devices, directly improving system performance
  - Maintains previously enabled CXL 3.x protocol enhancements with the 256B Flit format
  - Introduces the concept of native x2 width to support increased fan-out in the platform
  - Support for up to four retimers for increased channel reach
- CXL bundled port
  - Ability to aggregate device ports between Host and CXL accelerators (Type 1/2 devices) to increase bandwidth of the connection
- Memory RAS enhancements



# CXL 4.0: Doubles Bandwidth with Same Latency

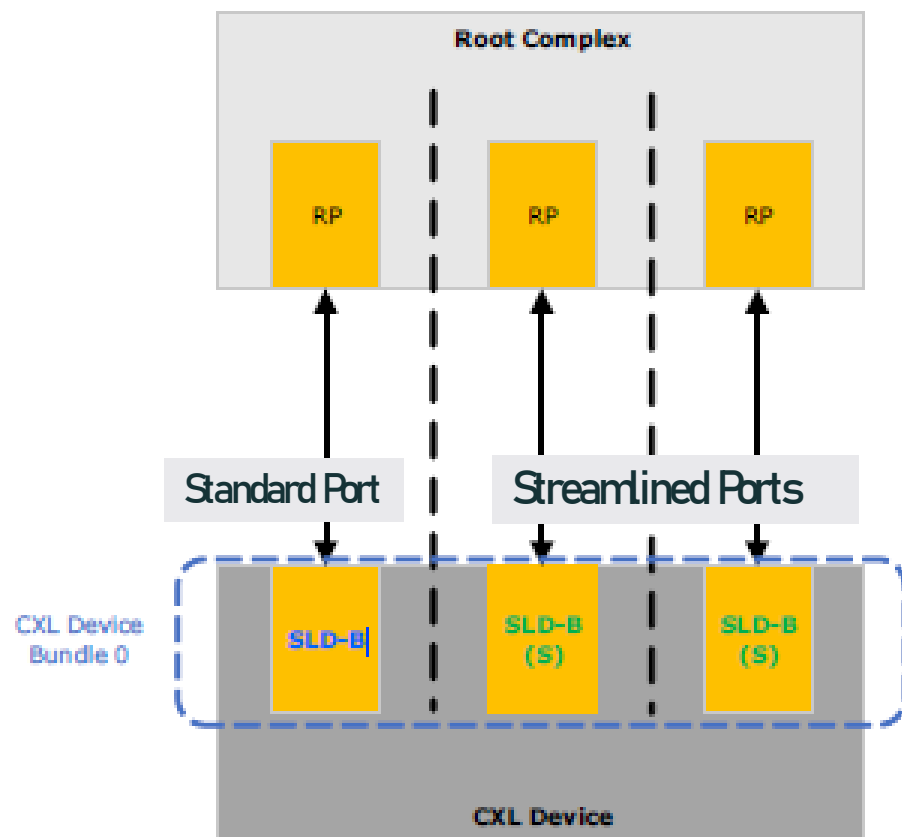


- Uses PCIe® 7.0 PHY @ 128 GT/s
- PCIe 7.0 FEC and CRC
  - No changes from CXL 3.0 FEC and CRC
  - Optical support with PCIe 7.0
  - Up to 4 Retimers for channel extension
- Standard 256B Flit along with an additional 256B Latency Optimized Flit (0-latency adder over CXL 2.0 and CXL 3.X)
  - 0-latency adder trades off FIT (failure in time,  $10^9$  hours) from  $5 \times 10^{-8}$  to 0.026 and Link efficiency impact from 0.94 to 0.92 for 2-5ns latency savings (x16 - x4)
- Native x2 width support
- Extends to lower data rates (8G, 16G, 32G, 64G)
- Keeps several previously enabled CXL 3.X protocol enhancements with the 256B Flit format



1: D. Das Sharma, "A Low-Latency and Low-Power Approach for Coherency and Memory Protocols on PCI Express 6.0 PHY at 64.0 GT/s with PAM-4 Signaling", IEEE Micro, Mar/ Apr 2022 (<https://ieeexplore.ieee.org/document/9662217>)

# Bundled Ports to Increased Bandwidth

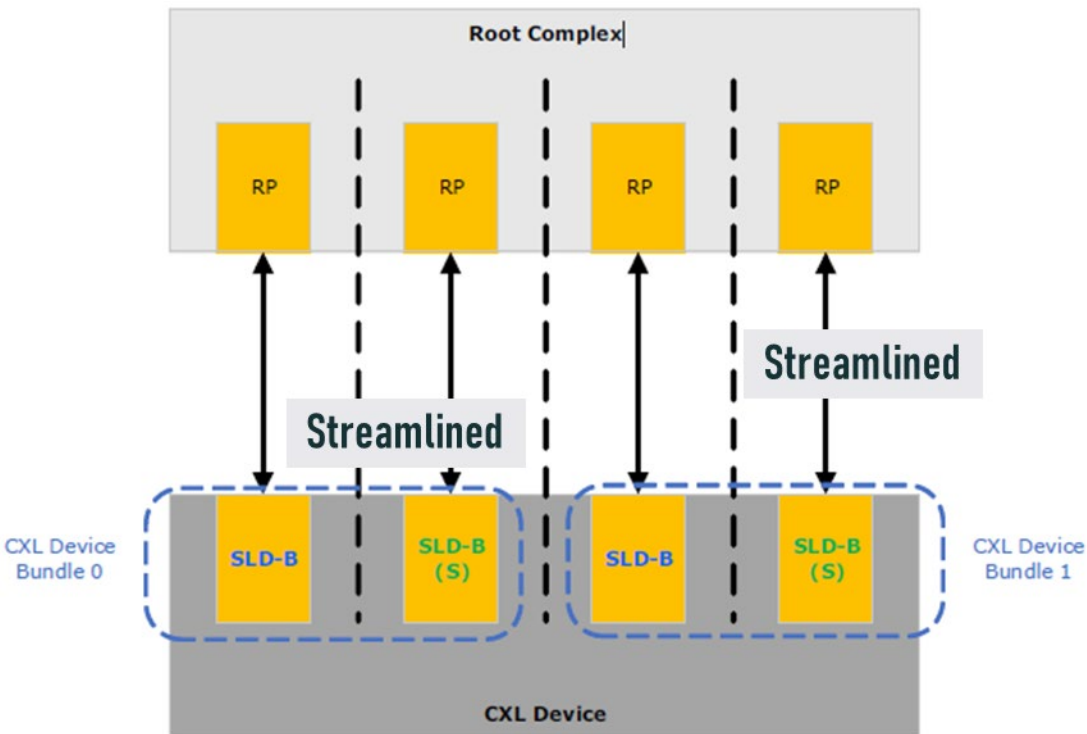


SLD-B : Single Logical Device  
exposed by each port in a bundle

- Bandwidth requirements growing, but at different rates depending on type of devices and workloads
- Certain devices and workloads can benefit from more than 2X Bandwidth scaling that transition to 128 GT/s provides
- Need the ability to logical aggregate multiple CXL Ports
- MH-SLDs enable such aggregation for memory expansion use case, but does not address accelerators
- CXL 4.0 introduces “Bundled Ports” => enables Logical Aggregation of multiple CXL Ports of an accelerator device
  - Type 1 and Type 2 Devices
  - Type 3 accelerator devices

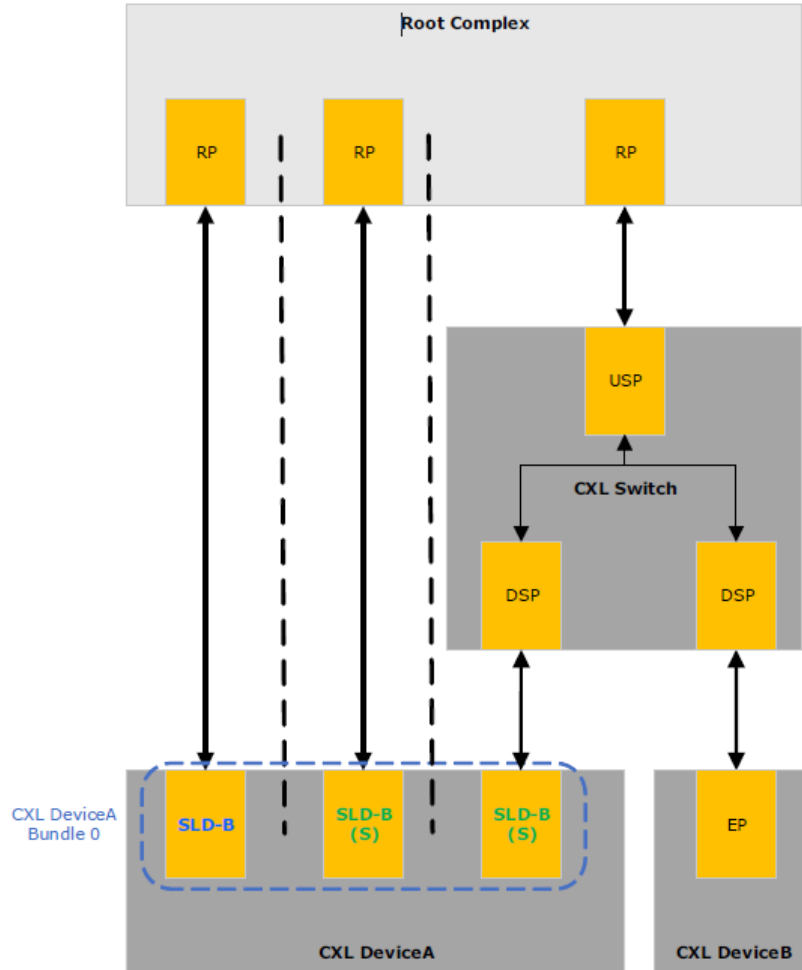
Challenge 5: Heterogeneous workloads demand higher bandwidth than ever

# Bundled Port Device (BPD) Construction

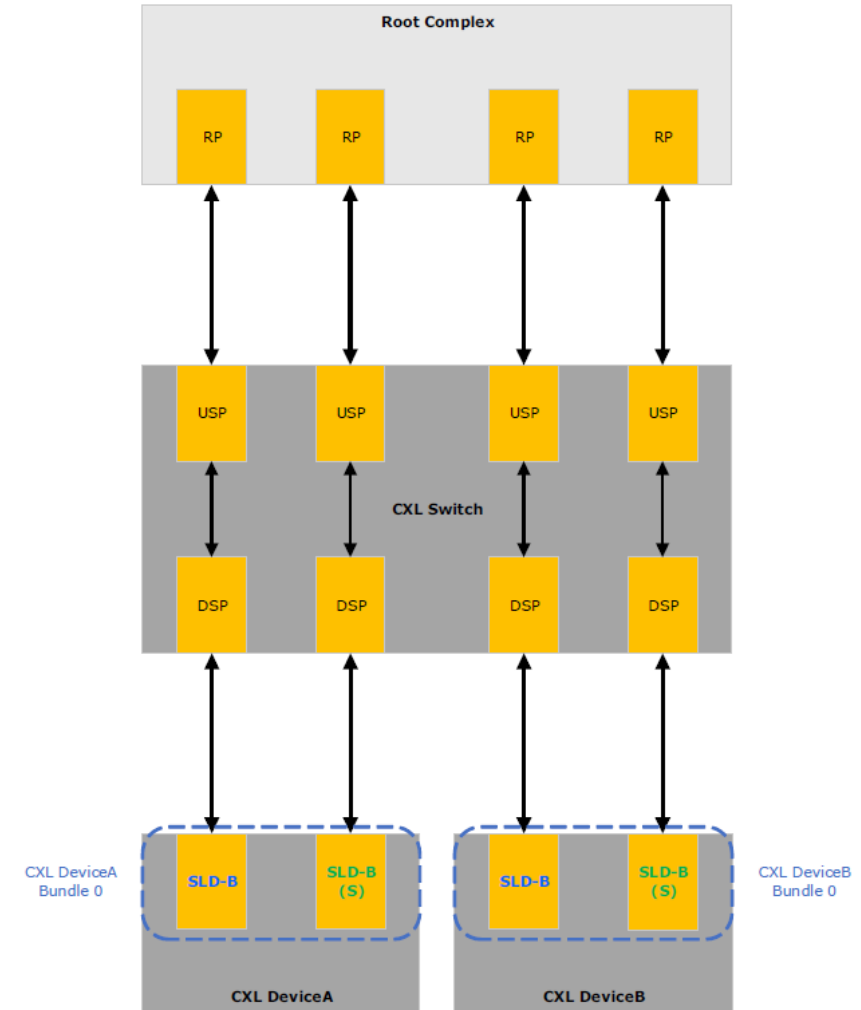


- One or more Bundle(s) per Device
- Each Bundle exposes at least one standard full-capability Port and any number of Streamlined Ports
- Streamlined Ports are area/power optimized for data bandwidth expansion
  - 256B Flit mode only
  - Optimized for UIO (non-UIO VC0 perf can be sub-optimal)
- Legacy software can safely enumerate BPD and manage the individual ports
- New software is needed to take advantage of Bundling (e.g. interleaving traffic across BPD ports)
- BPDs will typically use a different Device ID(s) to prevent non-BPD aware device drivers from managing the device
- Bundled Port-aware software is expected to configure IOMMU instances associated with the Bundled Ports so that all ports have an identical view of the memory
  - One port of BPD may issue ATS request and other port may utilize the returned HPA

# Switch Topologies - I

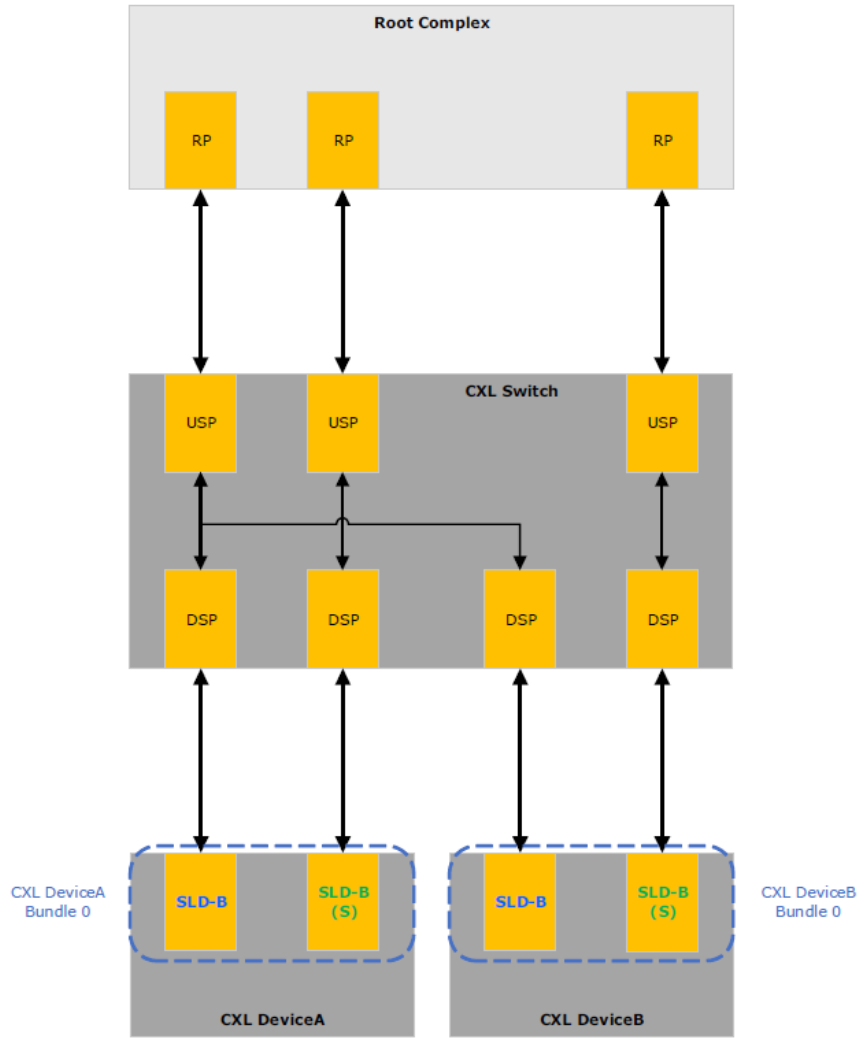


Bundled Port Connected to Different Entities

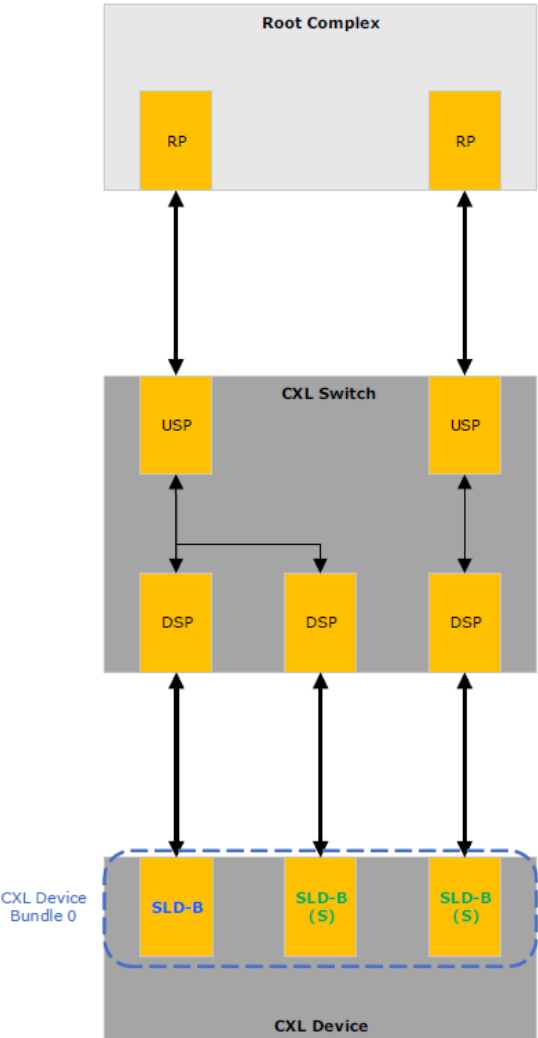


1:1 Port Mapping

# Switch Topologies - II



Dedicated Streamlined Port Paths



Merged Paths

- For PM/Reset purposes, each Port in a bundle is independent
  - Each port returns CDAT independently. HDM capacity of a BLD is the sum of its SLD-Bs.
  - All BLD links shall support PM VDM exchange including GPF
  - Each link can independently observe hot-resets
- CXL Reset and Cache disable bit are implemented by a single SLD-B. These have global scope and affect the entire BLD.
- Both TSP and IDE controls are centralized through one port



- Advanced CVME enhancements adding granularity control and event generation for Patrol Scrub cycles
  - Benefits: Allows for general media event record to be populated with error count flag conditions for advanced system decisions
- Defines mechanism for Host-initiated Post Package Repair (PPR) maintenance operations
  - Benefits: Ensure reliable DRAM row repair by enabling device-initiated PPR at boot through a persistent configuration bit across resets
- Defines memory sparing maintenance operations at device boot and enables deferral to next boot
  - Benefits: Enable flexible memory repair at boot by supporting both device-initiated and host-deferred sparing operations for improved reliability and maintenance efficiency

# CXL Specification Feature Summary

Not Supported
✓ Supported

Features	CXL 1.0 / 1.1	CXL 2.0	CXL 3.x	CXL 4.0
Release date	2019	2020	2022-2024	2025
Max link rate	32GTs	32GTs	64GTs	128GTs
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)			✓	✓
Back invalidate capabilities on Type 3 devices (HDM-DB)			✓	✓
Fabric Manager API definition for PBR Switch			✓	✓
Host-to-Host communication with Global Integrated Memory (GIM) concept			✓	✓
Trusted-Execution-Environment (TEE) Security Protocol			✓	✓
Memory expander enhancements (up to 32-bit of meta data, RAS capability enhancements)			✓	✓
Security, compliance, and CXL Memory Device enhancements			✓	✓
CXL Bundled Port				✓
Memory RAS enhancements (granular event reporting, Post Package Repair (PPR), and flexible memory sparing operations)				✓

# Evolution of CXL

## CXL 4.0

- 128 GTs
- CXL Bundled Port
- Memory RAS enhancements

## CXL 3.x

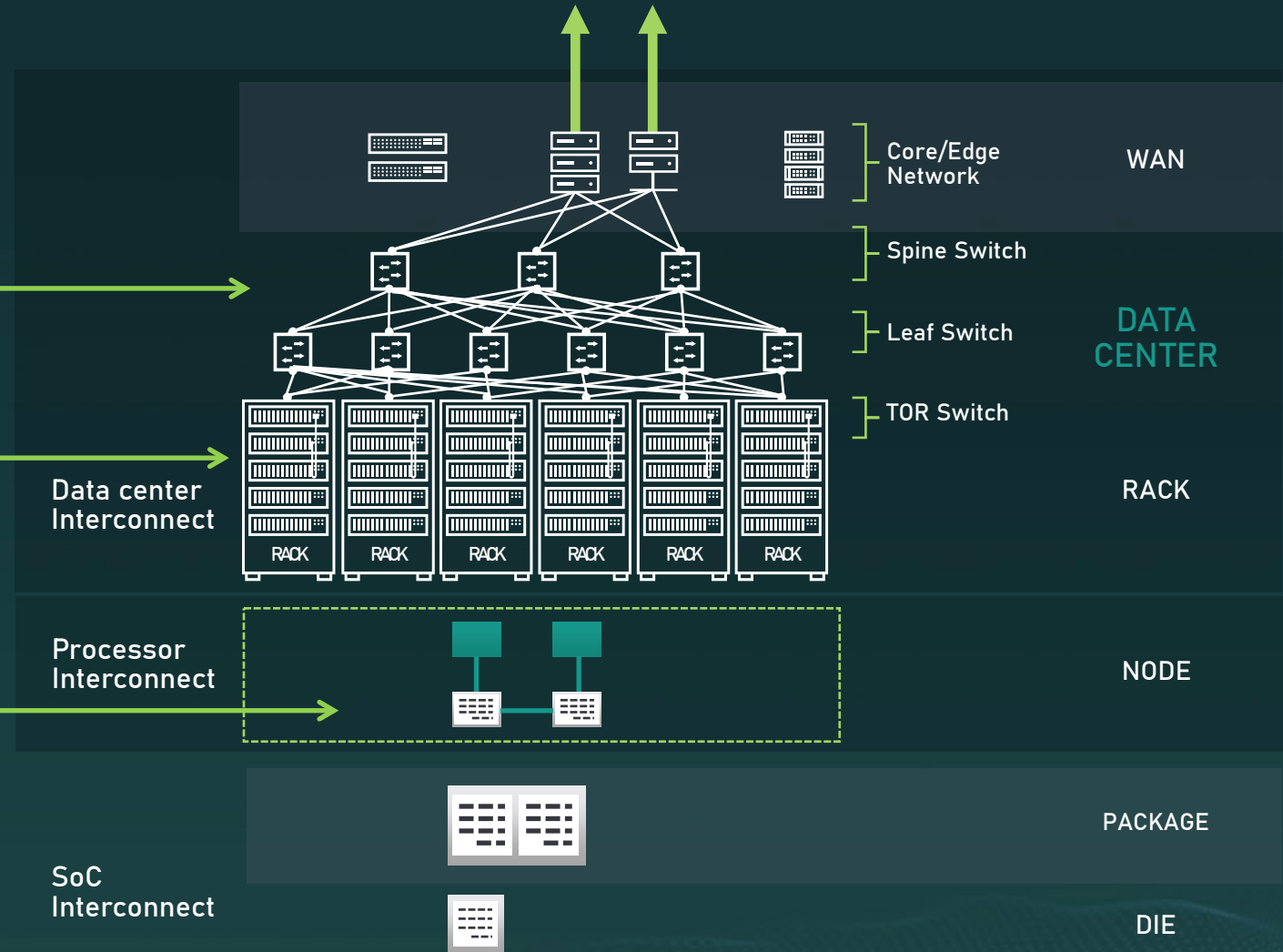
- 64 GTs
- Improved security
- Composable Fabric growth for disaggregation of memory & accelerators
- Memory sharing with back invalidate

## CXL 2.0

- Multiple nodes inside a Rack/Chassis supporting pooling of resources
- Switch-based memory pooling
- Global Persistent Flush (GPF)
- Link-level encryption

## CXL 1.1

- 32 GTs
- Single Node coherent interconnect
- Load/store semantics
- Hardware enforced cache coherence
- Low latency



# CXL: Health of the Ecosystem

Attribute	Status	Comments
Membership	280+ members	
Products	9 Compliance events since April 2023	9 <sup>th</sup> : : November 2-5, 2025. 30 CXL 1.1 and 30 CXL 2.0 devices in <a href="#">Integrators list</a> 7 Type-1, 8 Type-2, 57 Type-3, 6 Type1/2/3 Significant s/w development. <a href="#">Linux</a> Kernel 5.15 full support of T3 (Ubuntu 22.04.1 LTS/ Fedora Core 36 works) Multiple show-cases and demos in multiple conferences (SC, FMS, OCP, Memcon, etc.)
Heterogeneous Compute (Type1/2)	Deployed	<a href="#">UberNIC</a> : low-latency (1/2) and high throughput (>2.5x) <a href="#">VM Migration</a>
Memory (Type-3)	Deployed	Wide deployment. Both bandwidth and capacity expansion. <a href="#">Reduces loaded latency</a> . Multiple media (DRAM and storage covered)
Pooling (CXL 2.0)	PoCs look promising	VM Elastic Memory demand: Pond showed 9% DRAM savings initially (still substantial; paper in ASPLOS 23) –likely to go up - direct attach. Data base elastic memory demand: SAP and Intel: works well for TPCC (negligible performance degradation even with switches). <a href="#">See paper</a> .
Sharing/ Fabric	WIP (CXL3+)	CXL 3 silicon development in progress. S/W: Work actively continues on CXL 3.x (e.g., a patchset to layer a filesystem on top of shared memory)

- CXL 4.0 increases speed and bandwidth to meet the increasing demands of emerging workloads placed on today's data centers.
  - Doubles the bandwidth to 128GTs with zero added latency
    - Introduces the concept of native x2 width to support increased fan-out in the platform
  - CXL bundled port
  - Memory RAS enhancements
  - Maintains backward compatibility with CXL 3.x, 2.0, 1.1, and 1.0
- CXL Consortium Technical Working Groups continues to evolve to meet the future usage models

# Questions?





Thank You

[www.ComputeExpressLink.org](http://www.ComputeExpressLink.org)