OPENFABRICS ALLIANCE

2021 OFA Virtual Workshop

UNDERSTANDING COMPUTE EXPRESS LINK 2.0: A CACHE-COHERENT INTERCONNECT

Jim Pappas, Chairman

CXL Consortium

CXL Compute Express Link™

INDUSTRY LANDSCAPE

- **Industry trends are driving demand for faster data processing and next-generation data center performance**

Proliferation of
Cloud Computing

Growth of
AI & Analytics

Cloudification of
the Network & Edge

CXL CONSORTIUM

CXL Board of Directors

Alibaba Group, AMD, arm, Cisco, Dell EMC, Facebook, Google, Huawei, Hewlett Packard Enterprise, IBM, Intel, Microchip, Microsoft, Xilinx

Industry Open Standard for
High Speed Communications

150+ Member Companies

CXL DELIVERS THE RIGHT FEATURES AND ARCHITECTURE

Challenges

Industry trends driving demand for faster data processing and next-gen data center performance

Increasing demand for heterogeneous computing and server disaggregation

Need for increased memory capacity and bandwidth

Lack of open industry standard to address next-gen interconnect challenges

CXL

An open industry-supported cache-coherent interconnect for processors, memory expansion and accelerators

Coherent Interface

Leverages PCIe with 3 mix-and-match protocols

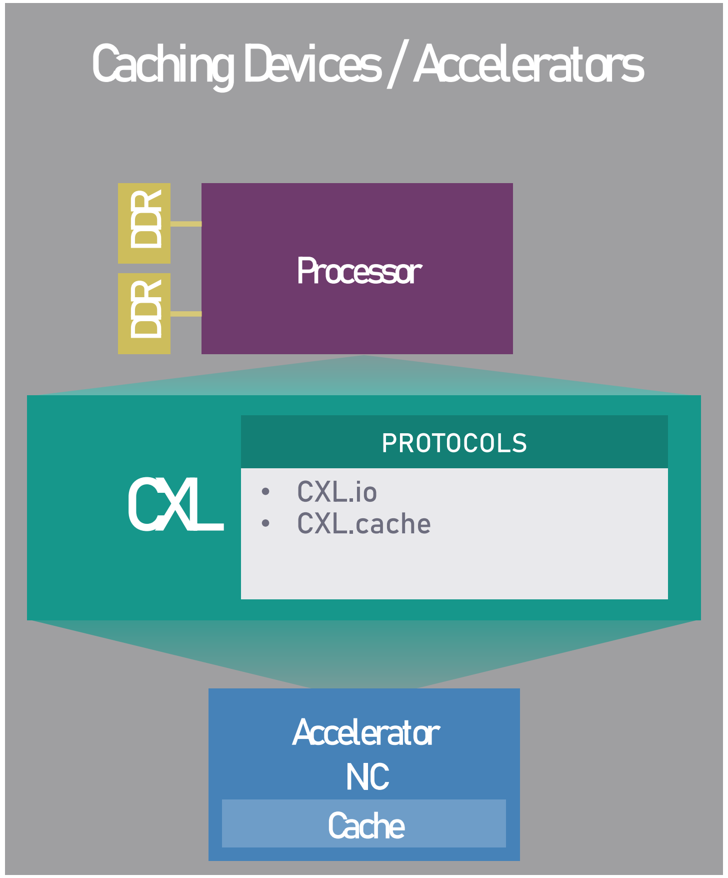
Low Latency

.Cache and .Memory targeted at near CPU cache coherent latency

Asymmetric Complexity

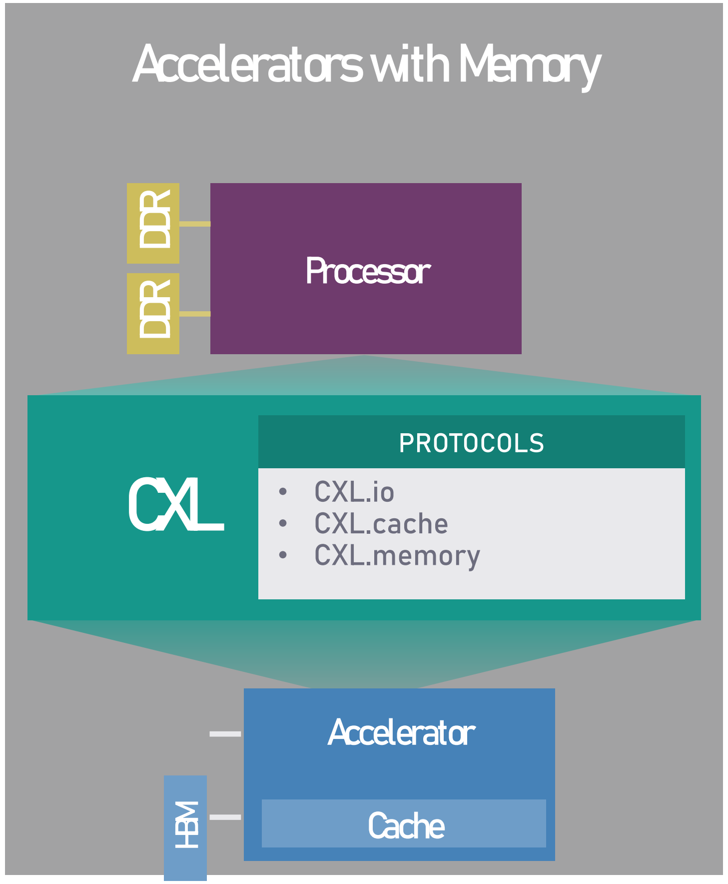
Eases burdens of cache coherent interface designs

REPRESENTATIVE CXL USAGES



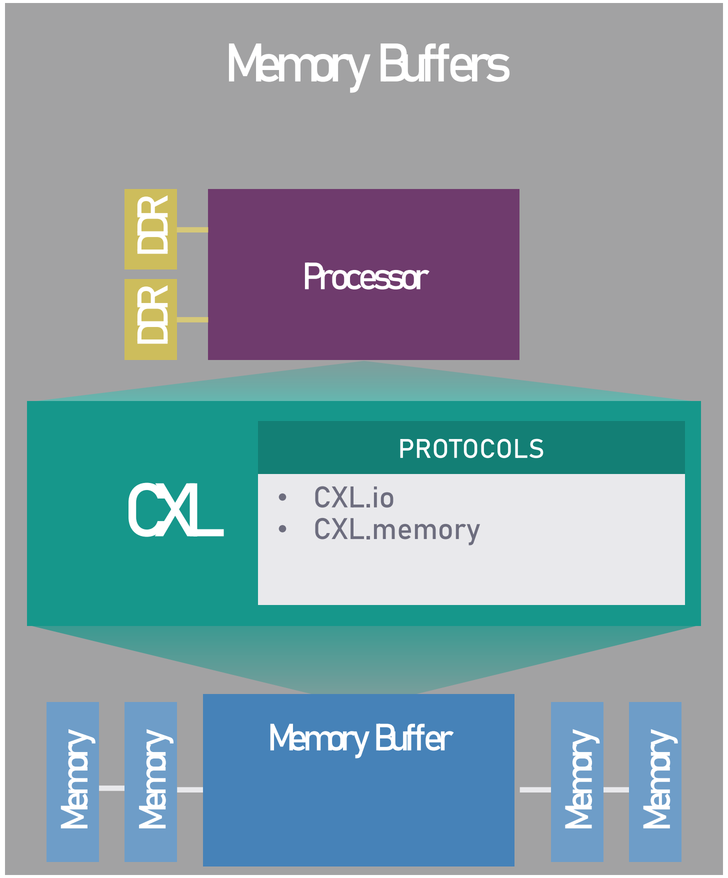
USAGES

- PGAS NIC
- NIC atomics



USAGES

- GP GPU
- Dense computation

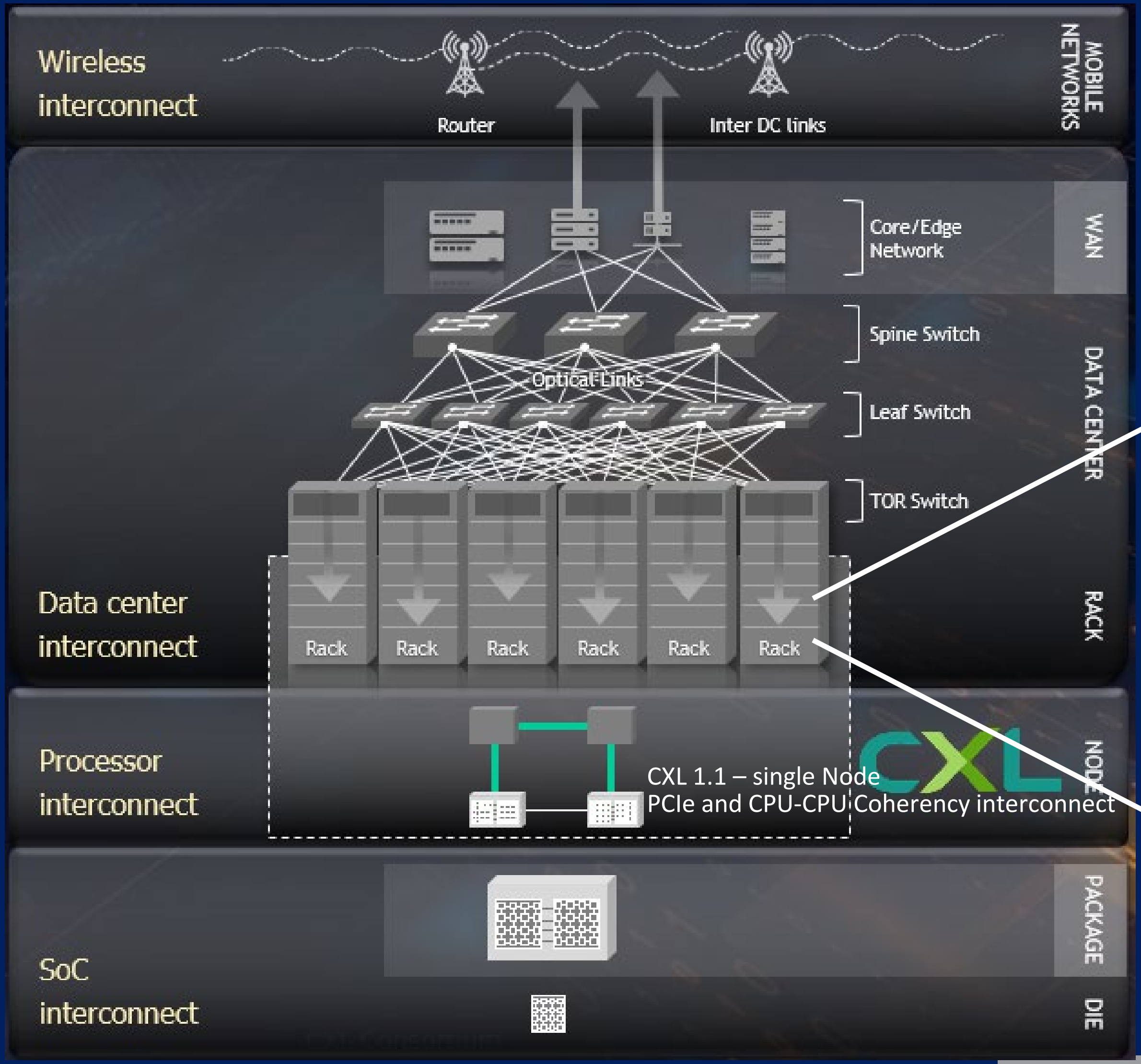


USAGES

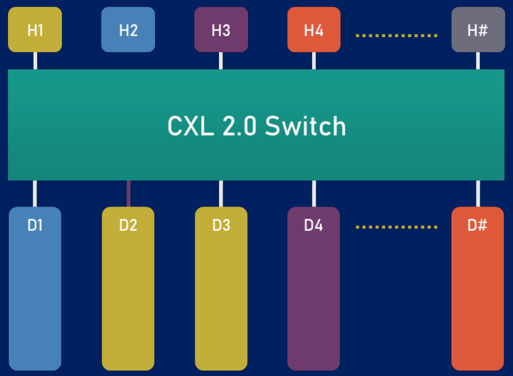
- Memory BW expansion
- Memory capacity expansion
- Storage class memory

THE NEW FEATURES OF CXL 2.0

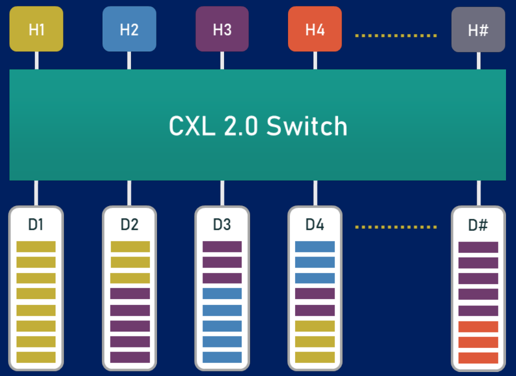
DATA CENTER: LOOKING OUTSIDE IN: SCOPE OF CXL 2.0 OVER CXL 1.1



Memory/Accelerator Pooling with Single Logical Devices



Memory Pooling with Multiple Logical Devices



CXL 2.0 Scope: Hot-Plug, Persistence, Switching, and Dis-aggregation

Feature	Description
CXL PCIe End-Point	CXL device to be discovered as PCIe Endpoint Support of CXL 1.1 devices directly connected to Root-Port or Downstream Switch Port
Switching	Single level of switching with multiple Virtual Hierarchies (cascaded possible in a single hierarchy) CXL Memory Fan-Out & Pooling with Interleaving CXL.Cache is direct routed between CPU and device with a single caching device within a hierarchy. Downstream port must be capable of being PCIe.
Resource Pooling	Memory Pooling for Type3 device – Multiple Logical Device (MLD), a single device to be pooled across 16 Virtual Hierarchies.
CXL.cache and CXL.mem enhancements	Persistence (Global Persistence Flush), Managed Hot-Plug, Function Level Reset Scope Clarification, Enhanced FLR for CXL Cache/Mem, Memory Error Reporting and QoS Telemetry
Security	Authentication and Encryption – CXL.IO uses PCIe IDE, CXL defines similar capability for CXL.$Mem
Software Infrastructure/ API	ACPI & UEFI ECNs to cover notification and management of CXL Ports and devices CXL Switch API for a multi-host or memory pooled CXL switch configuration and management

CXL 2.0 is fully backwards compatible with CXL 1.0/1.1 (see next slide for details)
CXL 2.0 spec Rev 0.7 in Q1, 2020; Rev 0.9 in Q2, 2020, and CXL 2.0 in Q3, 2020
Predictable spec release cadence by CXL consortium to help the ecosystem plan better.

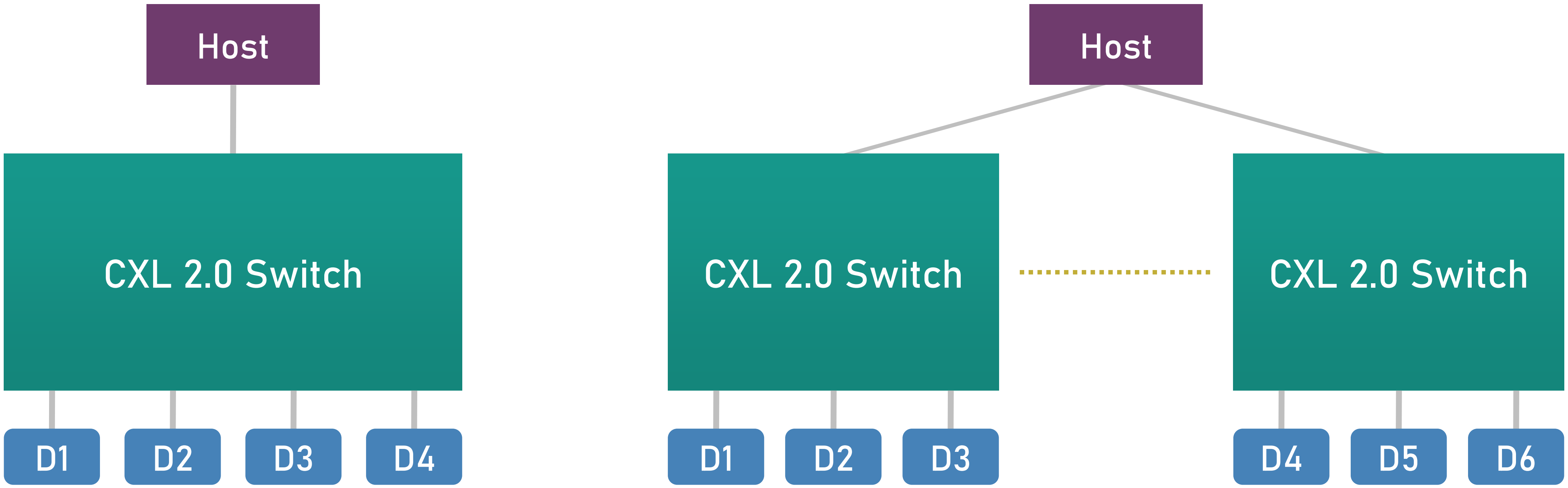
CXL 2.0 Scope: Hot-Plug, Persistence, Switching, and Dis-aggregation

Feature	Description
CXL PCIe End-Point	CXL device to be discovered as PCIe Endpoint Support of CXL 1.1 devices directly connected to Root-Port or Downstream Switch Port
Switching	Single level of switching with multiple Virtual Hierarchies (cascaded possible in a single hierarchy) CXL Memory Fan-Out & Pooling with Interleaving CXL.Cache is direct routed between CPU and device with a single caching device within a hierarchy. Downstream port must be capable of being PCIe.
Resource Pooling	Memory Pooling for Type3 device – Multiple Logical Device (MLD), a single device to be pooled across 16 Virtual Hierarchies.
CXL.cache and CXL.mem enhancements	Persistence (Global Persistence Flush), Managed Hot-Plug, Function Level Reset Scope Clarification, Enhanced FLR for CXL Cache/Mem, Memory Error Reporting and QoS Telemetry
Security	Authentication and Encryption – CXL.IO uses PCIe IDE, CXL defines similar capability for CXL.$Mem
Software Infrastructure/ API	ACPI & UEFI ECNs to cover notification and management of CXL Ports and devices CXL Switch API for a multi-host or memory pooled CXL switch configuration and management

CXL 2.0 is <u>fully backwards compatible</u> with CXL 1.0/1.1 (see next slide for details)
CXL 2.0 spec Rev 0.7 in Q1, 2020; Rev 0.9 in Q2, 2020, and CXL 2.0 in Q3, 2020
Predictable spec release cadence by CXL consortium to help the ecosystem plan better.

BENEFIT OF CXL 2.0 SWITCHING

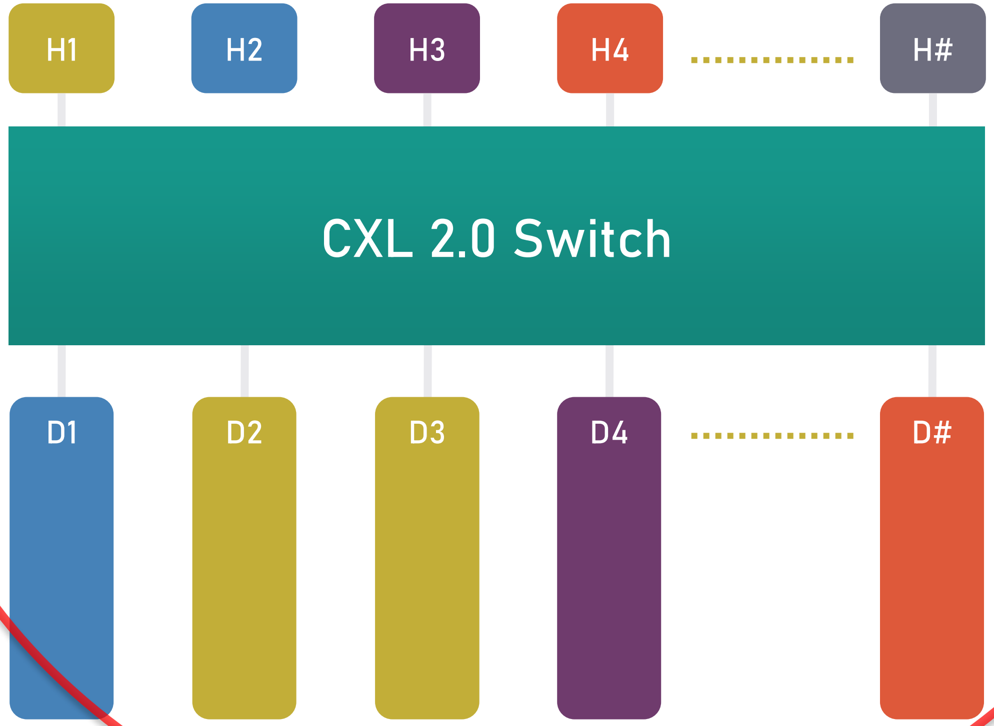
Expansion



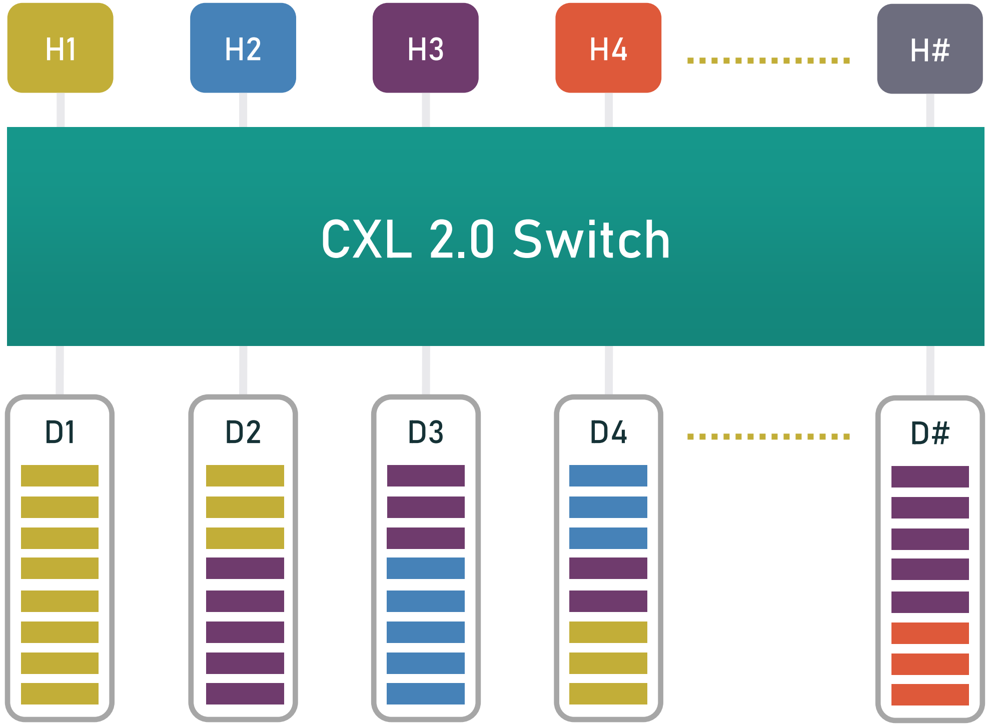
BENEFIT OF CXL 2.0 SWITCHING

Pooling

Memory/Accelerator Pooling with Single Logical Devices

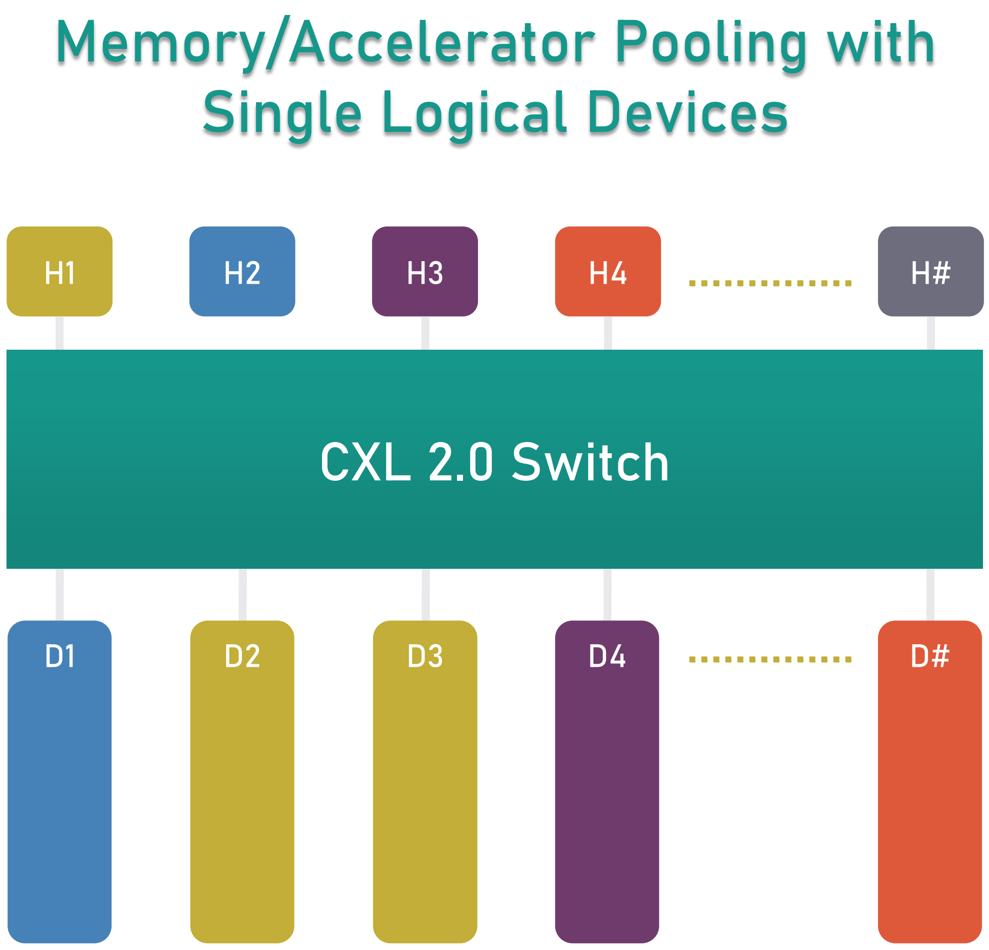


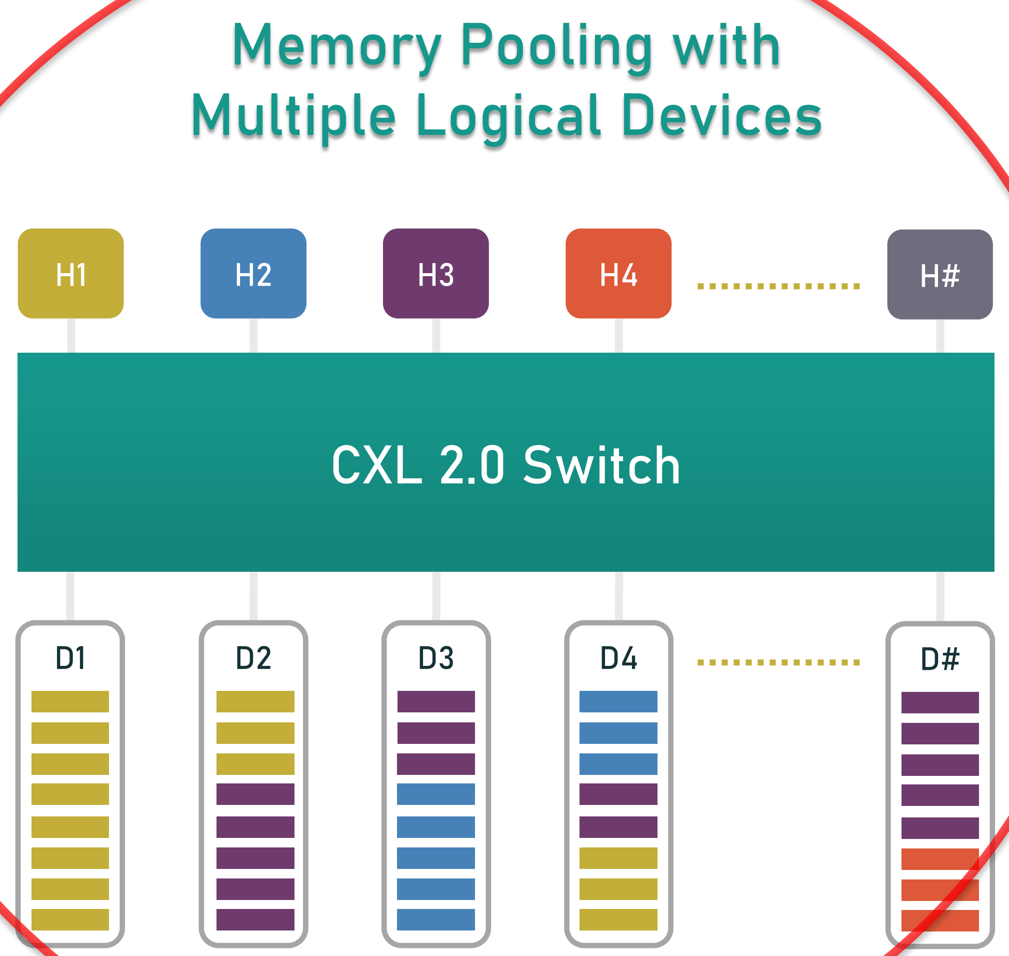
Memory Pooling with Multiple Logical Devices



BENEFIT OF CXL 2.0 SWITCHING

Pooling





CXL 2.0 Scope: Hot-Plug, Persistence, Switching, and Dis-aggregation

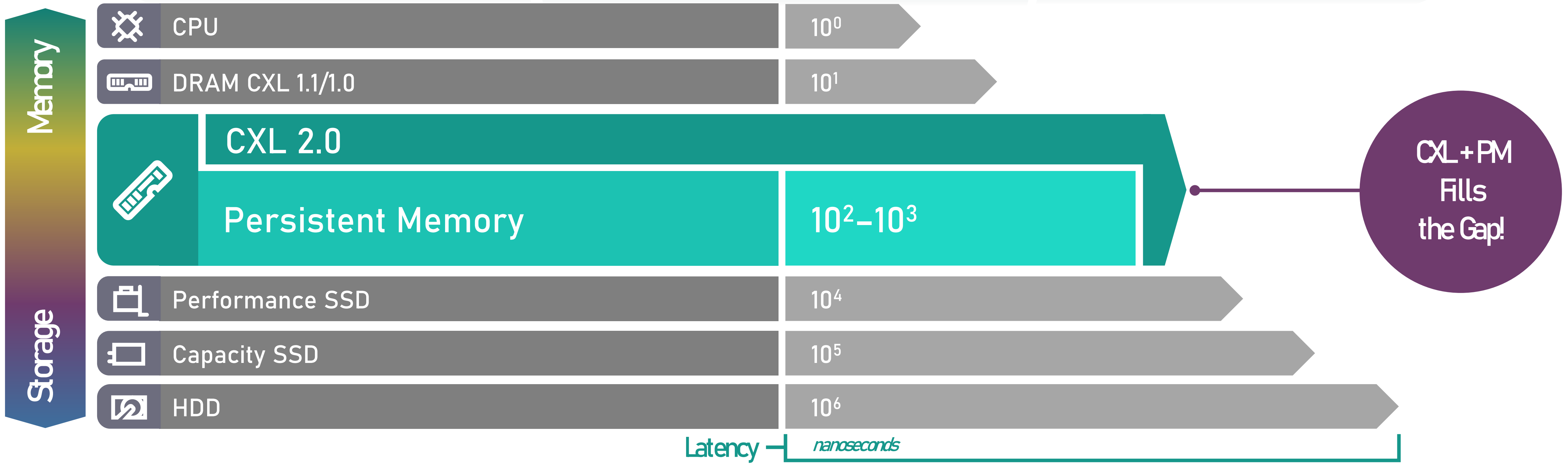
Feature	Description
CXL PCIe End-Point	CXL device to be discovered as PCIe Endpoint Support of CXL 1.1 devices directly connected to Root-Port or Downstream Switch Port
Switching	Single level of switching with multiple Virtual Hierarchies (cascaded possible in a single hierarchy) CXL Memory Fan-Out & Pooling with Interleaving CXL.Cache is direct routed between CPU and device with a single caching device within a hierarchy. Downstream port must be capable of being PCIe.
Resource Pooling	Memory Pooling for Type3 device – Multiple Logical Device (MLD), a single device to be pooled across 16 Virtual Hierarchies.
CXL.cache and CXL.mem enhancements	Persistence (Global Persistence Flush), Managed Hot-Plug, Function Level Reset Scope Clarification, Enhanced FLR for CXL Cache/Mem, Memory Error Reporting and QoS Telemetry
Security	Authentication and Encryption – CXL.IO uses PCIe IDE, CXL defines similar capability for CXL.$Mem
Software Infrastructure/ API	ACPI & UEFI ECNs to cover notification and management of CXL Ports and devices CXL Switch API for a multi-host or memory pooled CXL switch configuration and management

CXL 2.0 is <u>fully backwards compatible</u> with CXL 1.0/1.1 (see next slide for details)
CXL 2.0 spec Rev 0.7 in Q1, 2020; Rev 0.9 in Q2, 2020, and CXL 2.0 in Q3, 2020
Predictable spec release cadence by CXL consortium to help the ecosystem plan better.

BENEFITS OF CXL 2.0

and Persistent Memory

Moves Persistent Memory from Controller to CXL	Enables Standardized Management of the Memory and Interface	Supports a Wide Variety of Industry Form Factors



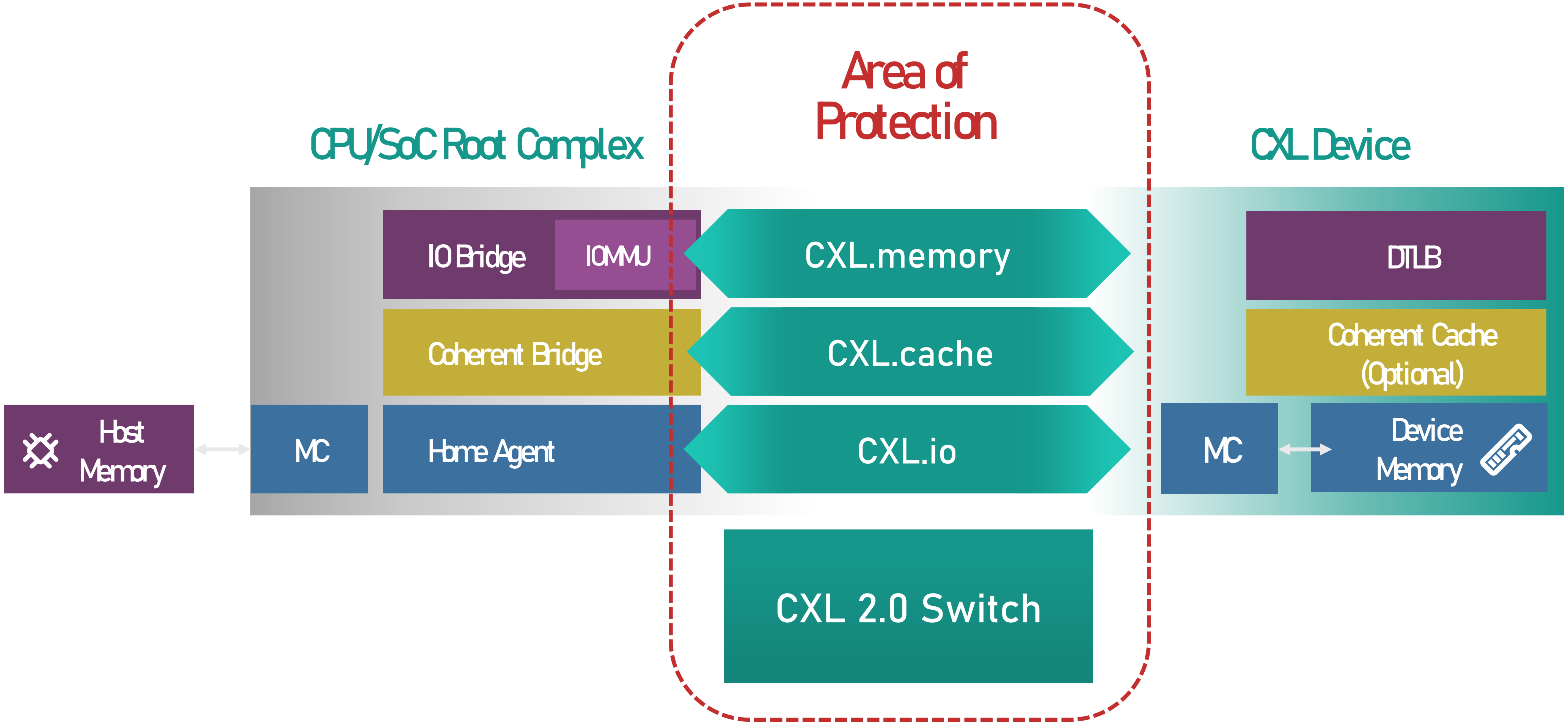
CXL 2.0 Scope: Hot-Plug, Persistence, Switching, and Dis-aggregation

Feature	Description
CXL PCIe End-Point	CXL device to be discovered as PCIe Endpoint Support of CXL 1.1 devices directly connected to Root-Port or Downstream Switch Port
Switching	Single level of switching with multiple Virtual Hierarchies (cascaded possible in a single hierarchy) CXL Memory Fan-Out & Pooling with Interleaving CXL.Cache is direct routed between CPU and device with a single caching device within a hierarchy. Downstream port must be capable of being PCIe.
Resource Pooling	Memory Pooling for Type3 device – Multiple Logical Device (MLD), a single device to be pooled across 16 Virtual Hierarchies.
CXL.cache and CXL.mem enhancements	Persistence (Global Persistence Flush), Managed Hot-Plug, Function Level Reset Scope Clarification, Enhanced FLR for CXL Cache/Mem, Memory Error Reporting and QoS Telemetry
Security	Authentication and Encryption – CXL.IO uses PCIe IDE, CXL defines similar capability for CXL.$Mem
Software Infrastructure/ API	ACPI & UEFI ECNs to cover notification and management of CXL Ports and devices CXL Switch API for a multi-host or memory pooled CXL switch configuration and management

CXL 2.0 is <u>fully backwards compatible</u> with CXL 1.0/1.1 (see next slide for details)
CXL 2.0 spec Rev 0.7 in Q1, 2020; Rev 0.9 in Q2, 2020, and CXL 2.0 in Q3, 2020
Predictable spec release cadence by CXL consortium to help the ecosystem plan better.

CXL 2.0 SECURITY BENEFITS

CXL 2.0 provides Integrity and Data Encryption of traffic across all entities (Root Complex, Switch, Device)



CXL 2.0: BACKWARDS COMPATIBLE WITH CXL 1.0/1.1

CPU – Device Connectivity	CXL 1.X (1.1/1.0) EP	CXL 2.0 EP	PCIe EP/ Switch	Comments
CPU with CXL 1.x	CXL 1.x	CXL 1.x	PCIe	CXL 2.0 EP needs to support both RCiEP and EP modes
CPU with CXL 2.0	CXL 1.x	CXL 2.0	PCIe	CXL 2.0 CPU also needs to be bi-modal for backwards compatibility

CXL 2.0 Switch Connectivity	Operation	Comments
Upstream: CPU	Only CXL 2.0	Since switch definition is with CXL 2.0, the platform requirement is upstream port be a CXL 2.0 CPU
Downstream: CXL Device	CXL 1.x or CXL 2.0	All downstream CXL ports work as CXL 1.x or CXL 2.0 – mix and match to the device's capability
Downstream: PCIe EP/ Switch	PCIe	Any CXL switch downstream Port must be able to support a PCIe hierarchy, either an EP or a PCIe switch but assigned to *one* domain
Downstream: CXL Switch	N/A	CXL 2.0 is defined only as a single level switch for multiple virtual hierarchies (no cascading of CXL switches)

IN SUMMARY

CXL Consortium momentum continues to grow

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

CXL 2.0 introduces new features & usage models

- Switching, pooling, persistent memory support, security
- Fully backward compatible with CXL 1.1 and 1.0
- Built in Compliance & Interop program

Call to action

- Join CXL Consortium
- Visit www.computeexpresslink.org for more information
- Follow us on Twitter and LinkedIn for more updates!

Compute Express Link™

2021 OFA Virtual Workshop

THANK YOU

Jim Pappas, CXL Chairman

CXL Consortium