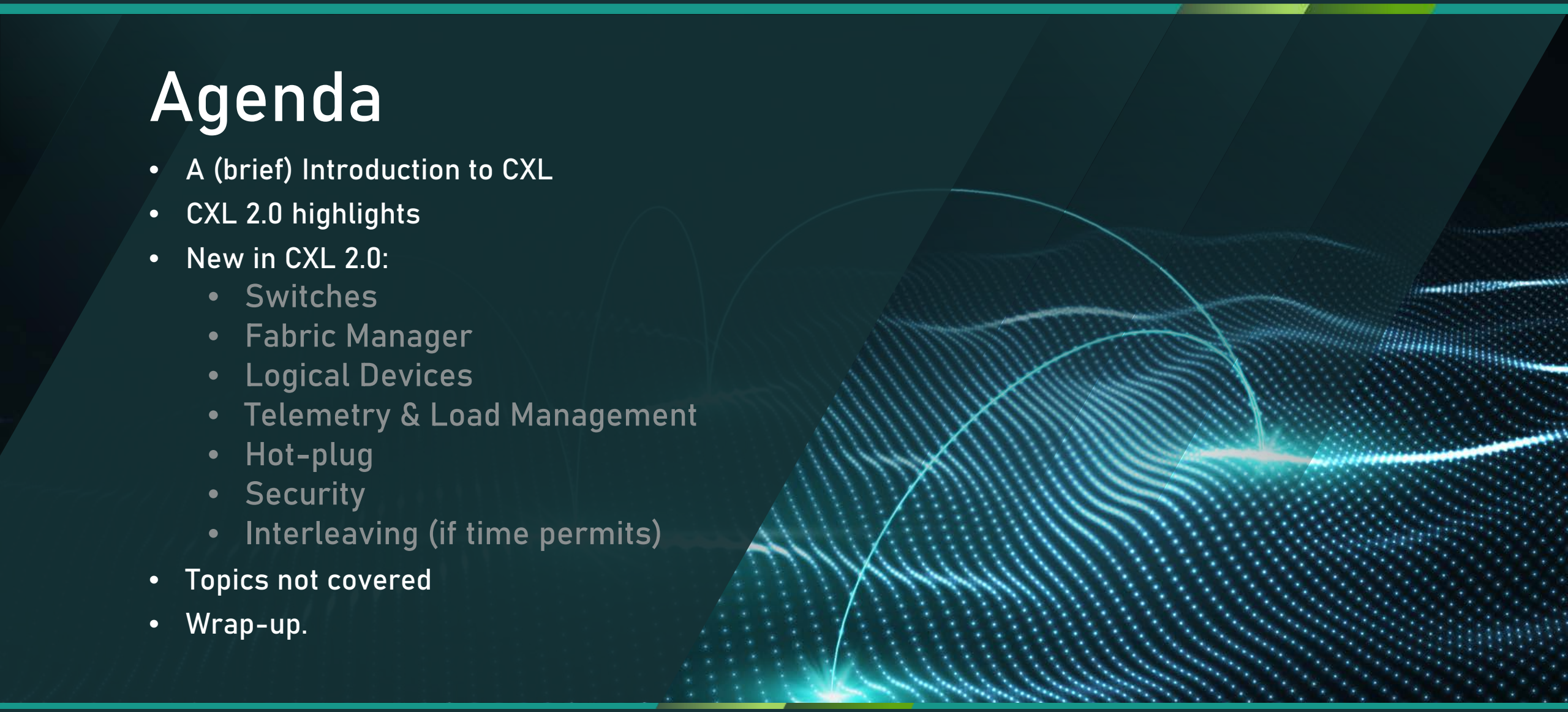
CXL 1.1 vs CXL 2.0 – What's the difference ?

Danny Volkind – CTO & Co-Founder, UnifabriX
Elad Shliselberg – System Architect, UnifabriX

June 15, 2022

Agenda

- A (brief) Introduction to CXL
- CXL 2.0 highlights
- New in CXL 2.0:
  - Switches
  - Fabric Manager
  - Logical Devices
  - Telemetry & Load Management
  - Hot-plug
  - Security
  - Interleaving (if time permits)
- Topics not covered
- Wrap-up.

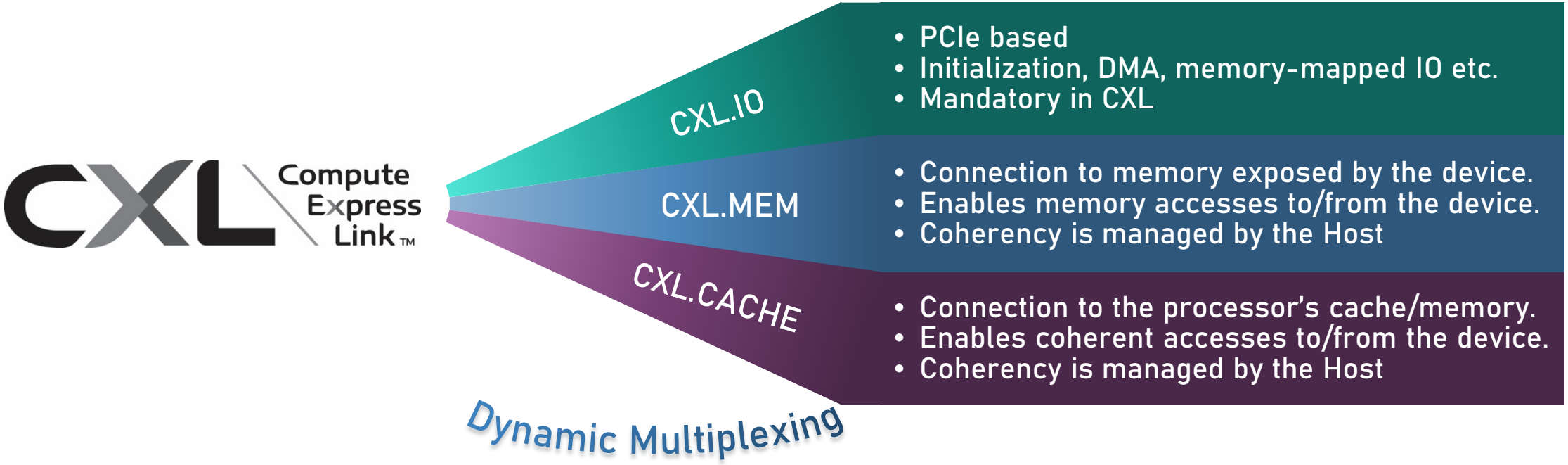


References

- Documents
  - Compute Express Link™ (CXL™) 1.1 Specification
  - CXL™ 2.0 Specification
  - CXL™ 2.0 ECN
  - PCI Express® (PCIe®) Base Specification Revision 5.0
  - Unified Extendible Firmware Interface (UEFI) Specification
- CXL Consortium Webinars

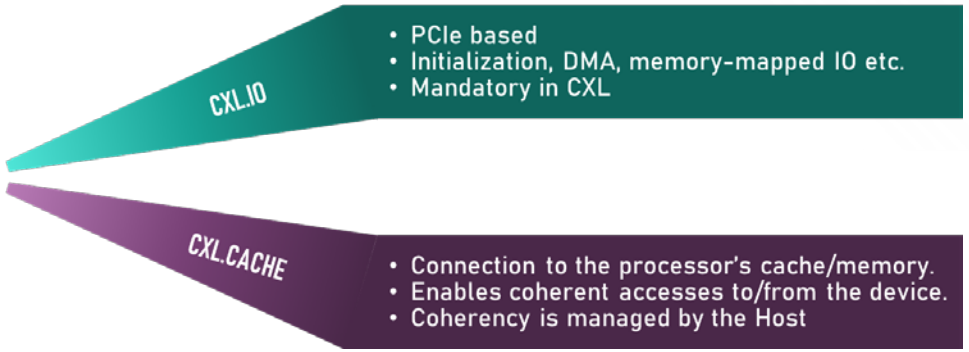
CXL – A Revolution in Computer Architecture

- A unified bus enabling a standard way of connecting to the IO, caching and memory sub-systems.
  - Wide industry support and adoption.
- Leverages PCI Express® (PCIe®)
  - Runs on top of a PCIe PHY (Gen5 down to Gen3 in degraded mode).
- Optimized for high-rates and low-latency.
- Comprised of three dynamically multiplexed protocols: CXL.IO, CXL.MEM, CXL.CACHE.



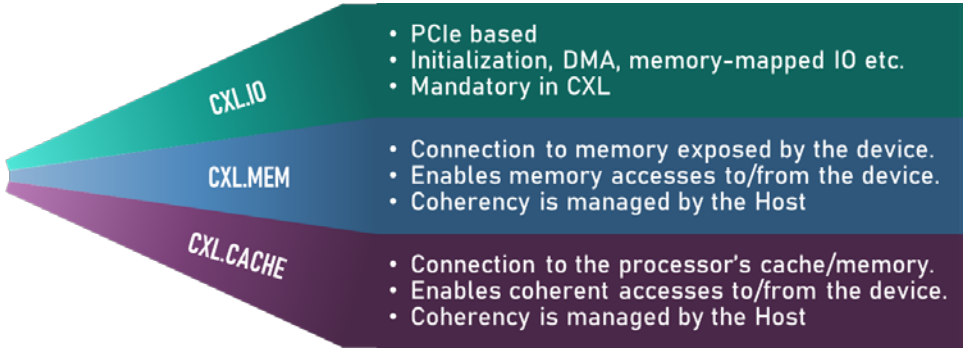
CXL Mx & Match Approach

Type 1 device



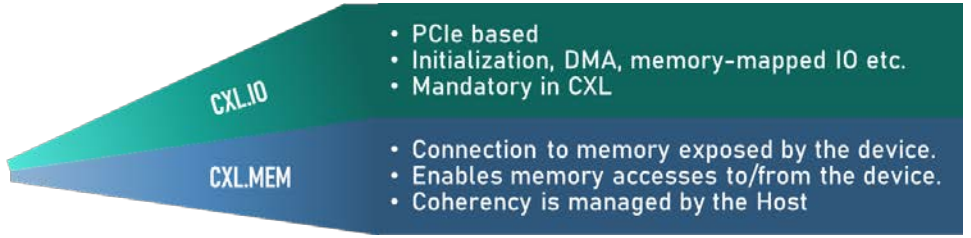
Caching/Accelerator device
An accelerator requiring coherent access to the processor's memory with no host access to its own memory.
- A smart NIC

Type 2 device



Accelerator device with memory
An accelerator requiring access to the processor's memory and host access to its own memory.
- GPU, FPGA, Computation accelerator

Type 3 device



Memory buffer/expander
Exposes memory resources to the host.
- Memory BW/capacity expansion

CXL Evolution

- Timeline
  - CXL 1.0 - March 2019
  - CXL 1.1 - Sep 2019
  - CXL 2.0 - Nov 2020
  - CXL 3.0 – in the making
- Wide support and adoption: Over 170 industry leading companies have joined the CXL consortium.
- Server CPU products supporting CXL 1.1 publicly announced and are expected by the end of this year.

How will it look like?

- A CXL device is detected during the PCIe training sequence exchange (APN mechanism).
- The platform FW produces the System Resource Affinity Table (SRAT) and Heterogeneous Memory Attribute Table (HMAT) – both defined in the UEFI specification.
- Provides the OS information about memory initiators and targets in the system and proximity domains.
- Also includes performance information such as latency and bandwidth

```
fedora kernel: acpi/hmat: HMAT: Memory Flags:0001 Processor Domain:0 Memory Domain:0
fedora kernel: acpi/hmat: HMAT: Memory Flags:0001 Processor Domain:1 Memory Domain:1
fedora kernel: acpi/hmat: HMAT: Memory Flags:0000 Processor Domain:0 Memory Domain:2
fedora kernel: acpi/hmat: HMAT: Memory Flags:0000 Processor Domain:0 Memory Domain:3
```

```
devuser@zeus.unifabrix.com ~>lsmem -b --output RANGE,SIZE,NODE,ZONES
RANGE                                  SIZE NODE  ZONES
0x0000000000000000-0x000000007fffffff    2147483648    0   None
0x0000000100000000-0x000000107fffffff   66571993088    0 Normal
0x0000001080000000-0x000000247fffffff   85899345920    1 Normal
0x0000002480000000-0x000000a47fffffff  549755813888    2 Normal
0x000000a480000000-0x000008a47fffffff 8796093022208    3 Normal

Memory block size:          2147483648
Total online memory:     9500467658752
Total offline memory:                0
```

CXL 2.0 Highlights

- Backward compatibility with CXL 1.1
- Fanout enhancement: switching and pooling
- Fabric Management
- Fine-grain resource allocation
- Telemetry and load-management
- Hot-plug support
- Security
- And much more…

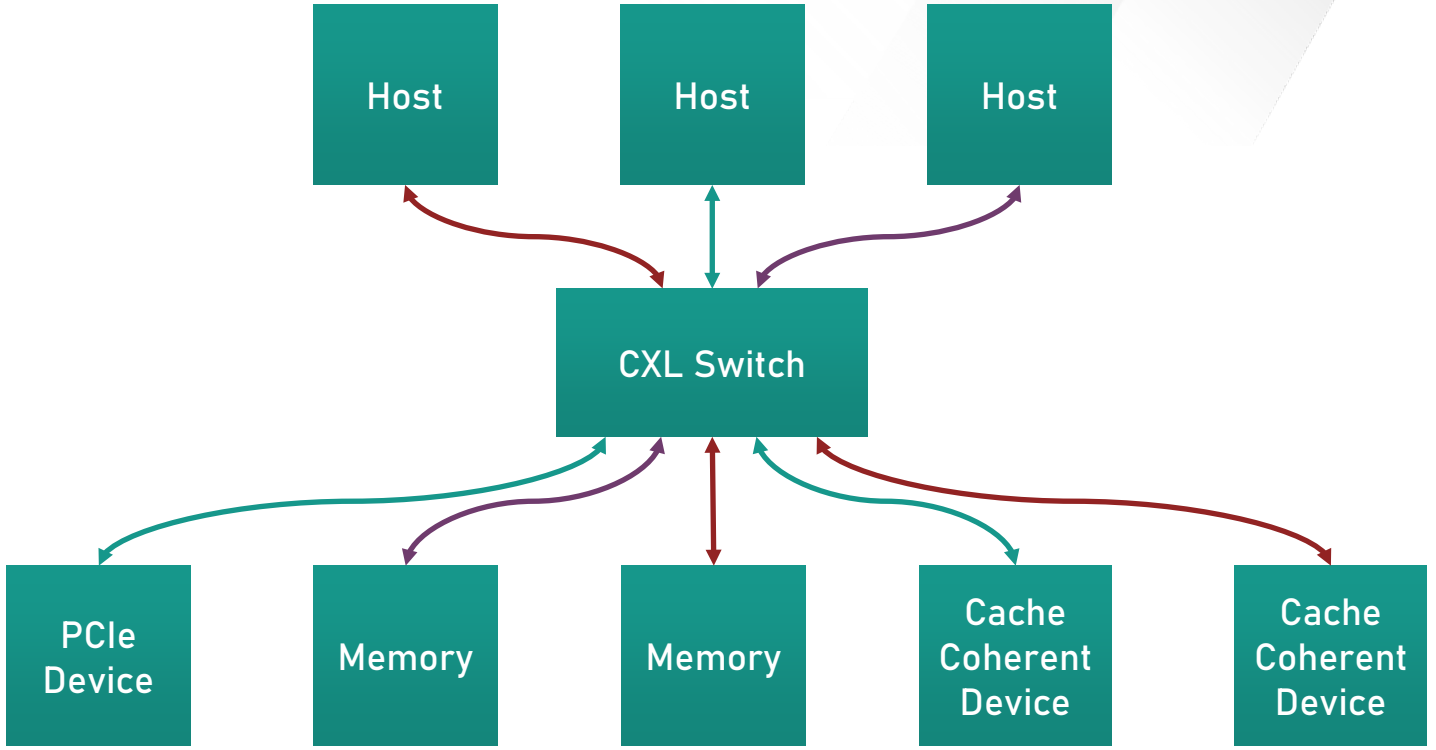
CXL 2.0

Facilitated Scaling Beyond the Platform

Management
Virtual Hierarchies
Resource Pooling
Security
Telemetry
Hot Plug

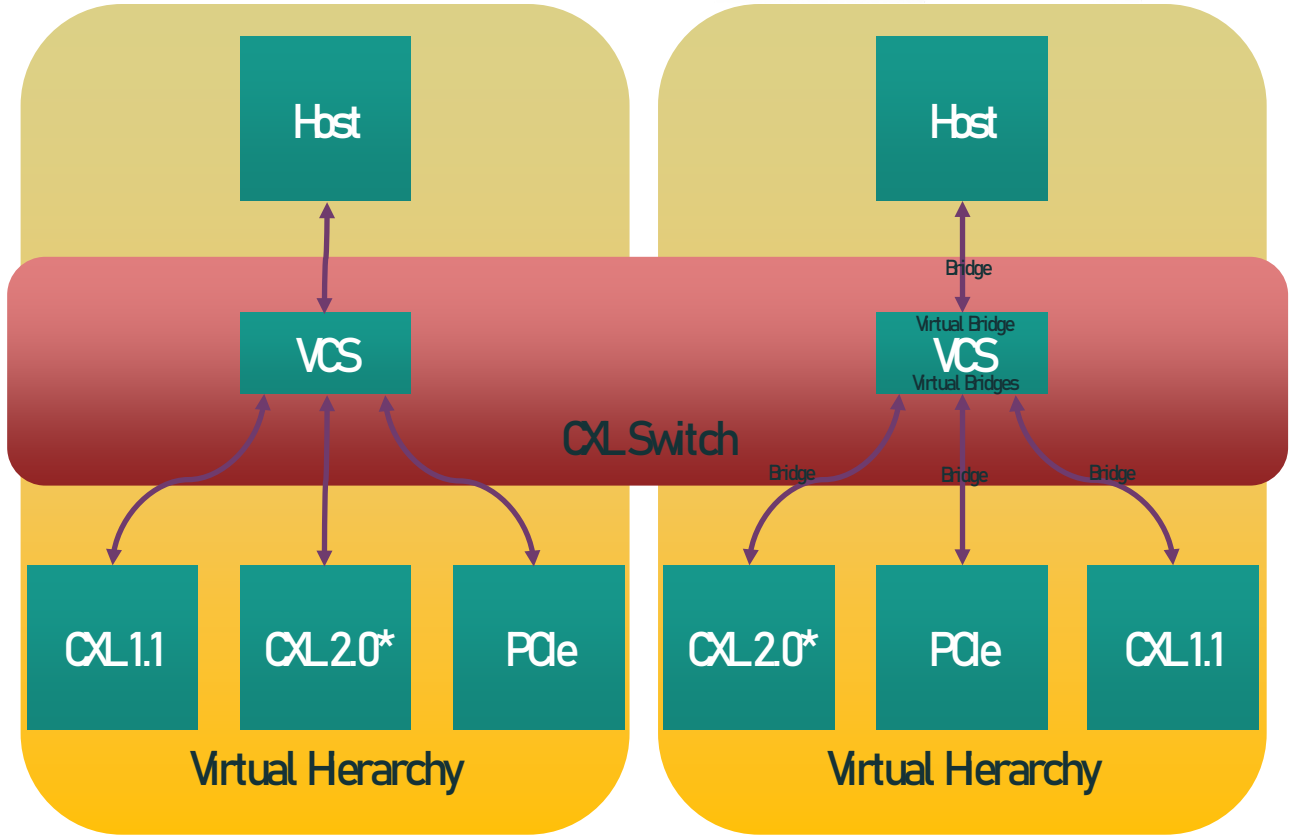
Switches

- What is a CXL Switch?
  - Multiple Host
  - Multiple Device
  - Single Layer
- Lynchpin of CXL 2.0



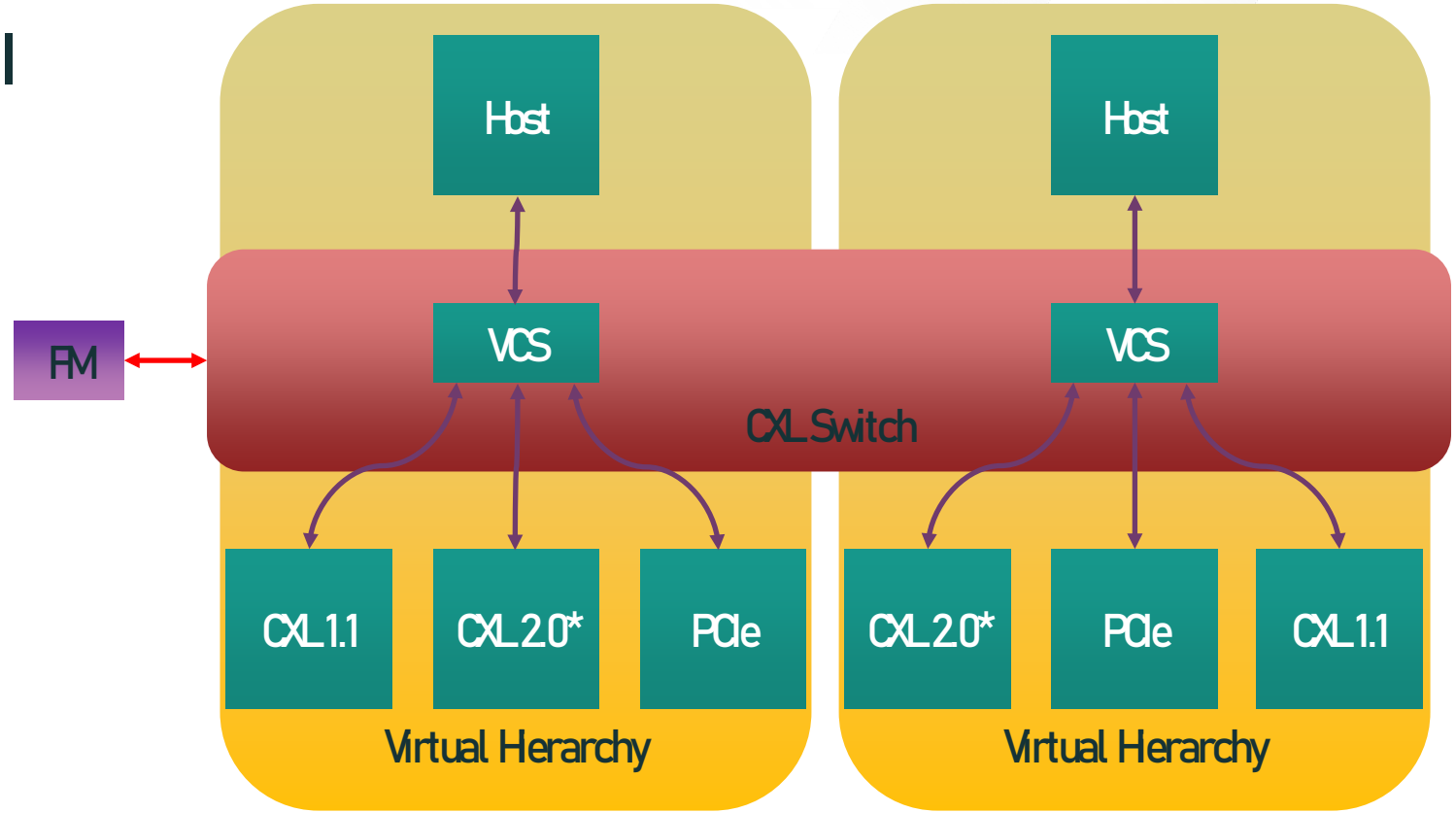
Switches – Hierarchies (VCS)

- Fine-grained resource allocation
- Virtual CXL Switch (VCS)
- Resource and host isolation



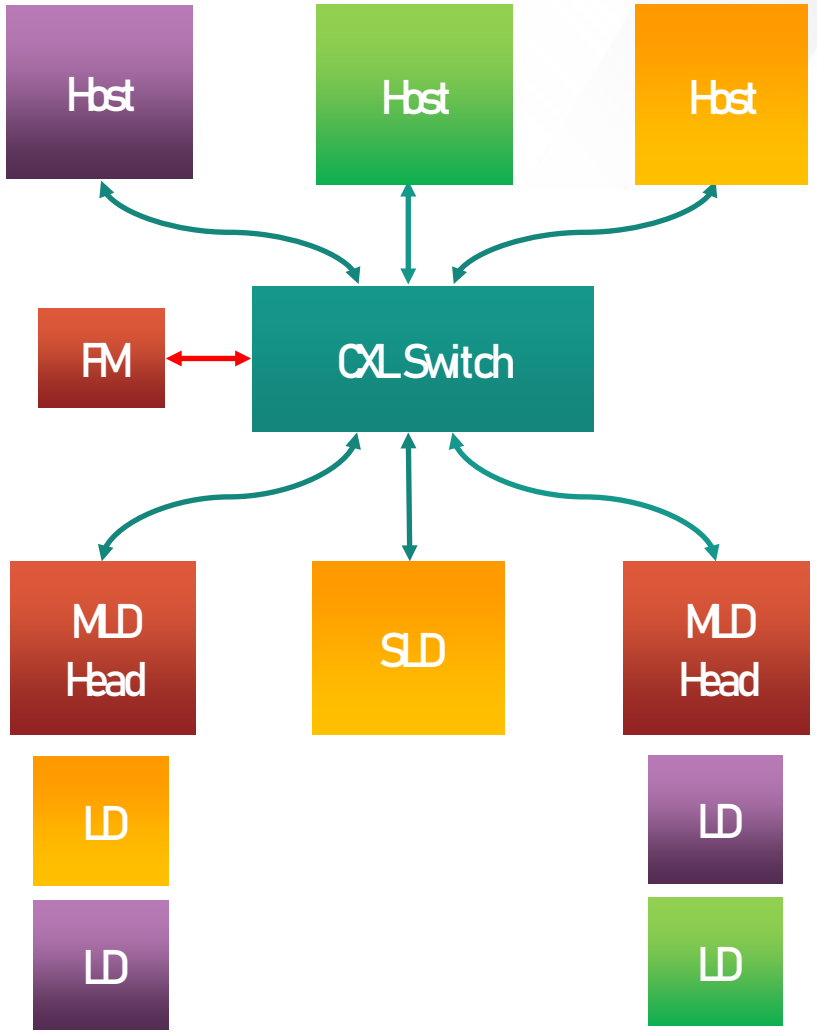
Fabric Management

- Central Manager
- Unified management API
- Enumerates System
- Centralized Control per switch
- Can be anything!
  - On board BMC
  - External controller
  - Centralized controller



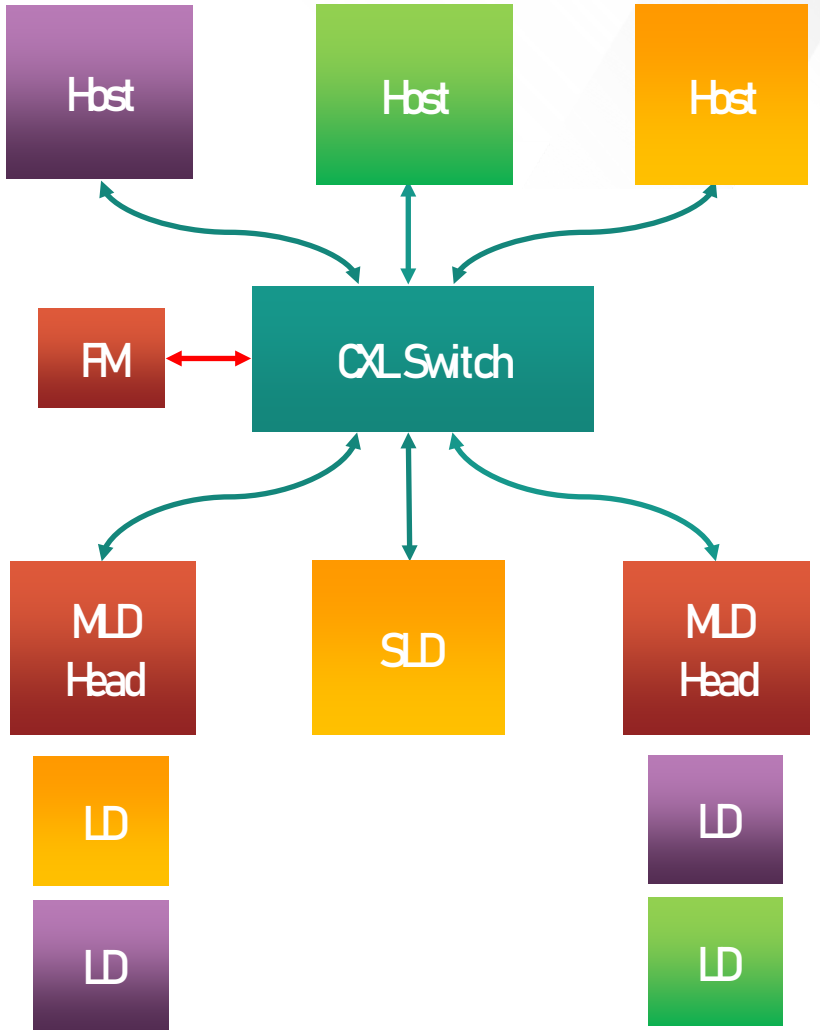
Logical Devices (LD)

- Multi-headed Device (1.1)
  - Multiple host connections
  - Single physical device
- Logical Device (LD)
  - Resides in single virtual hierarchy
  - Functional Interface
- Single Logical Device (SLD)
- What is an MLD?
  - Allows splitting and management of identifiable resources across multiple hosts



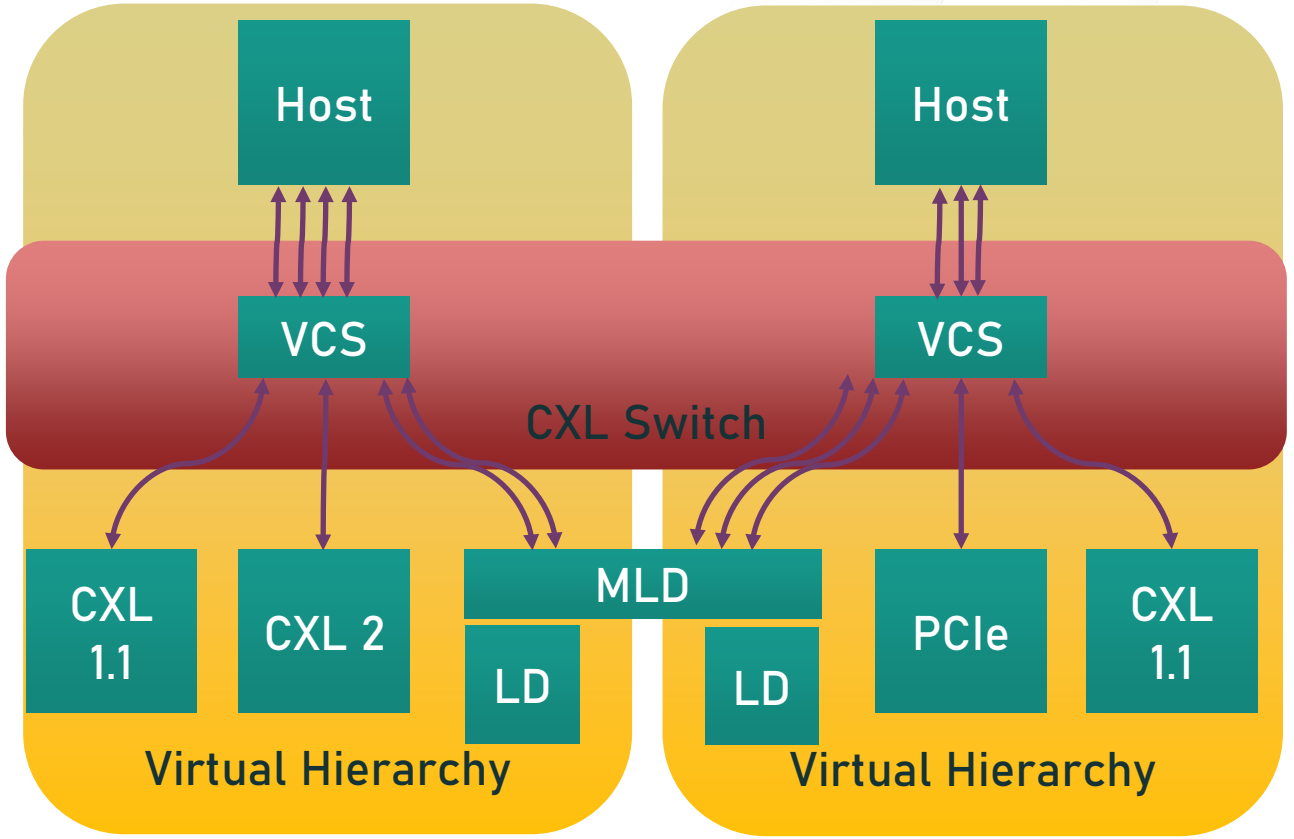
Multi-Logic Device (MLD)

- Switch oriented
  - Managed by Fabric Manager
- MLD can be split between virtual hierarchies
- Logical Device ID (LDID)
- Up to 16 T3 LDs per MLD



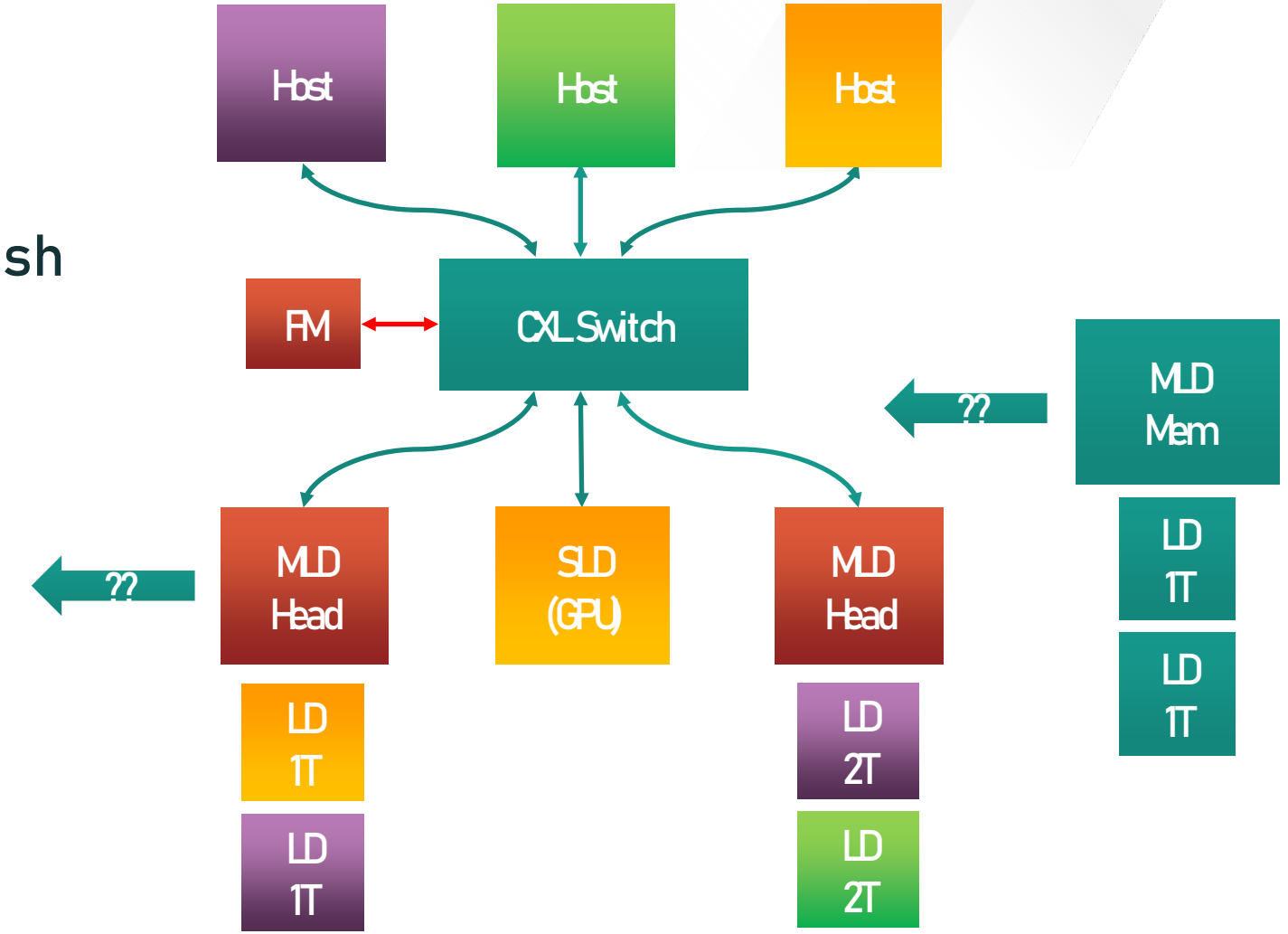
QoS Telemetry + Load Management

- Embedded in memory transactions
- Real-time load telemetry
- Run-time calibration



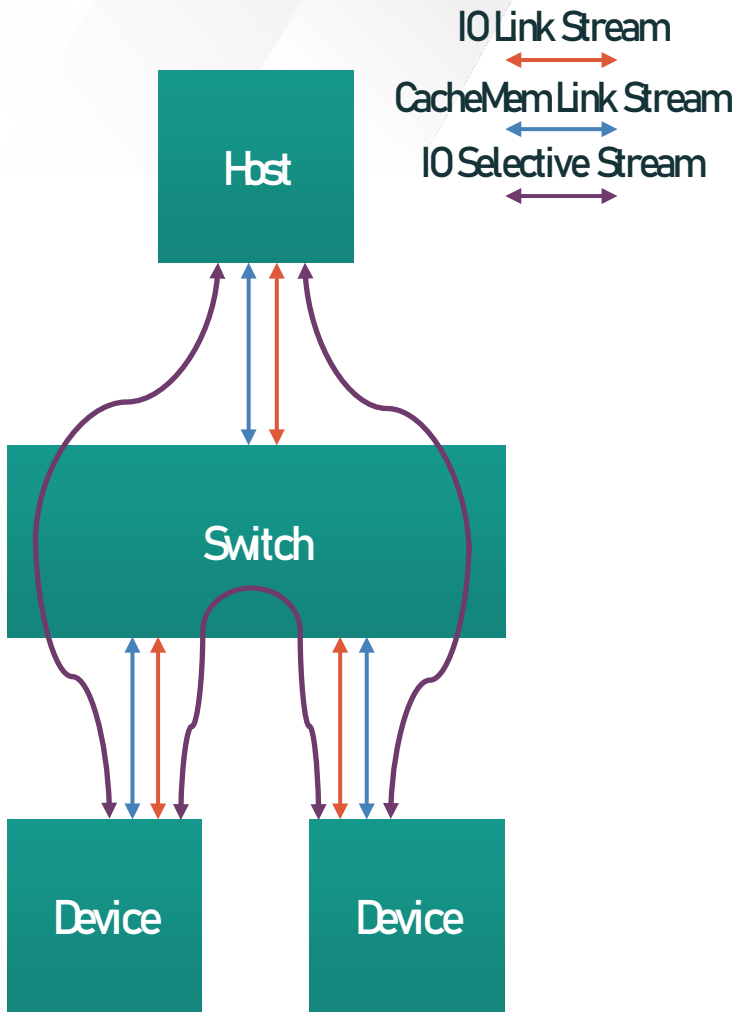
Hot Plug

- Hot Add
- Managed Hot Removal
  - GPF – Global Persistent Flush
- Management
  - Down Stream ports
  - Fabric Management
    - Allocation
    - MLDs



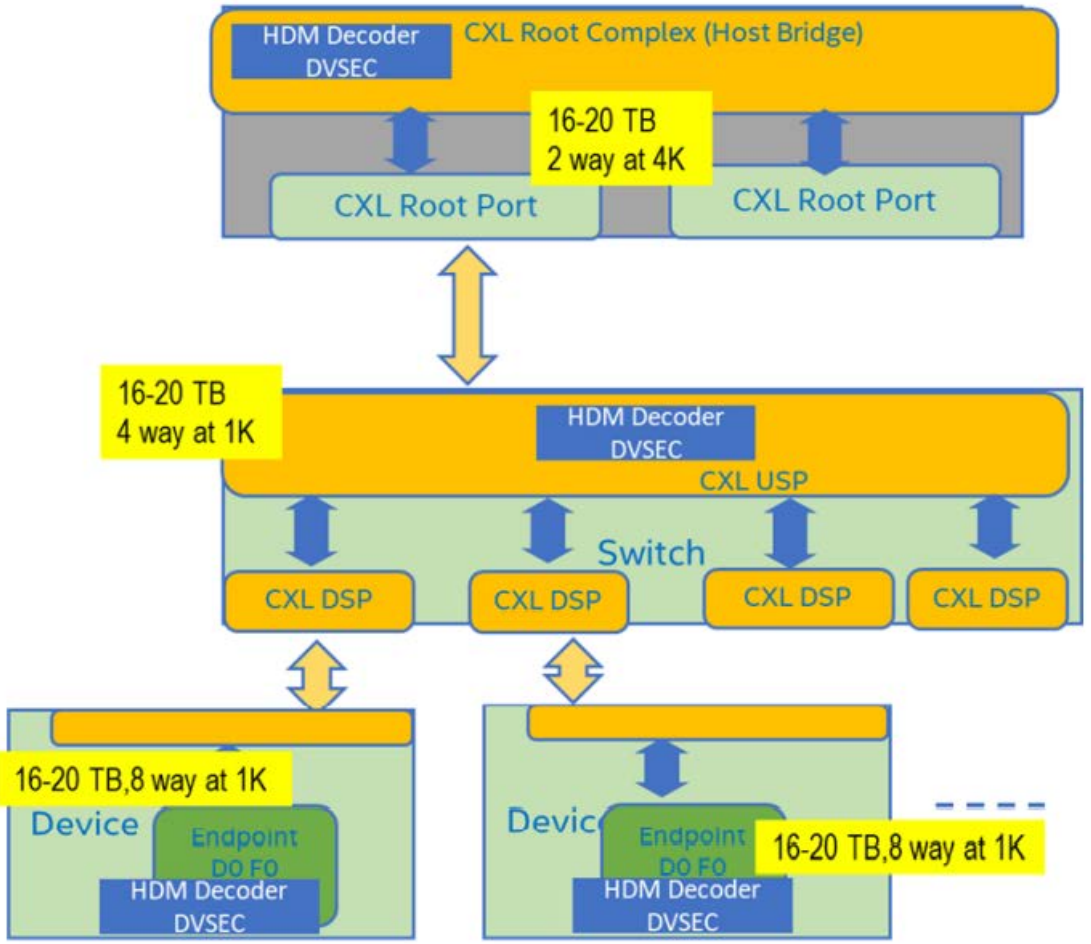
Security

- DMTF Security Protocol and Data Model (SPDM) PCIe IDE ECN
- Certificates that can be verified
- Asymmetric Handshake, Symmetric Encryption
  - EC-DSA
  - AES-GCM
    - Integrated MAC for data verification
    - Skid Mode for lower latency
- Link Level or Selective stream protection
  - CXL.IO vs CXL.CacheMem
  - Point to Point



Memory Interleaving

- Memory Interleaving groups
- Configurable
  - On address bits 14-8
  - granularity of 256B-16KB
  - Up to 16* way
  - Latency grouping
  - Cross Host Interleaving



Major Topics not Covered

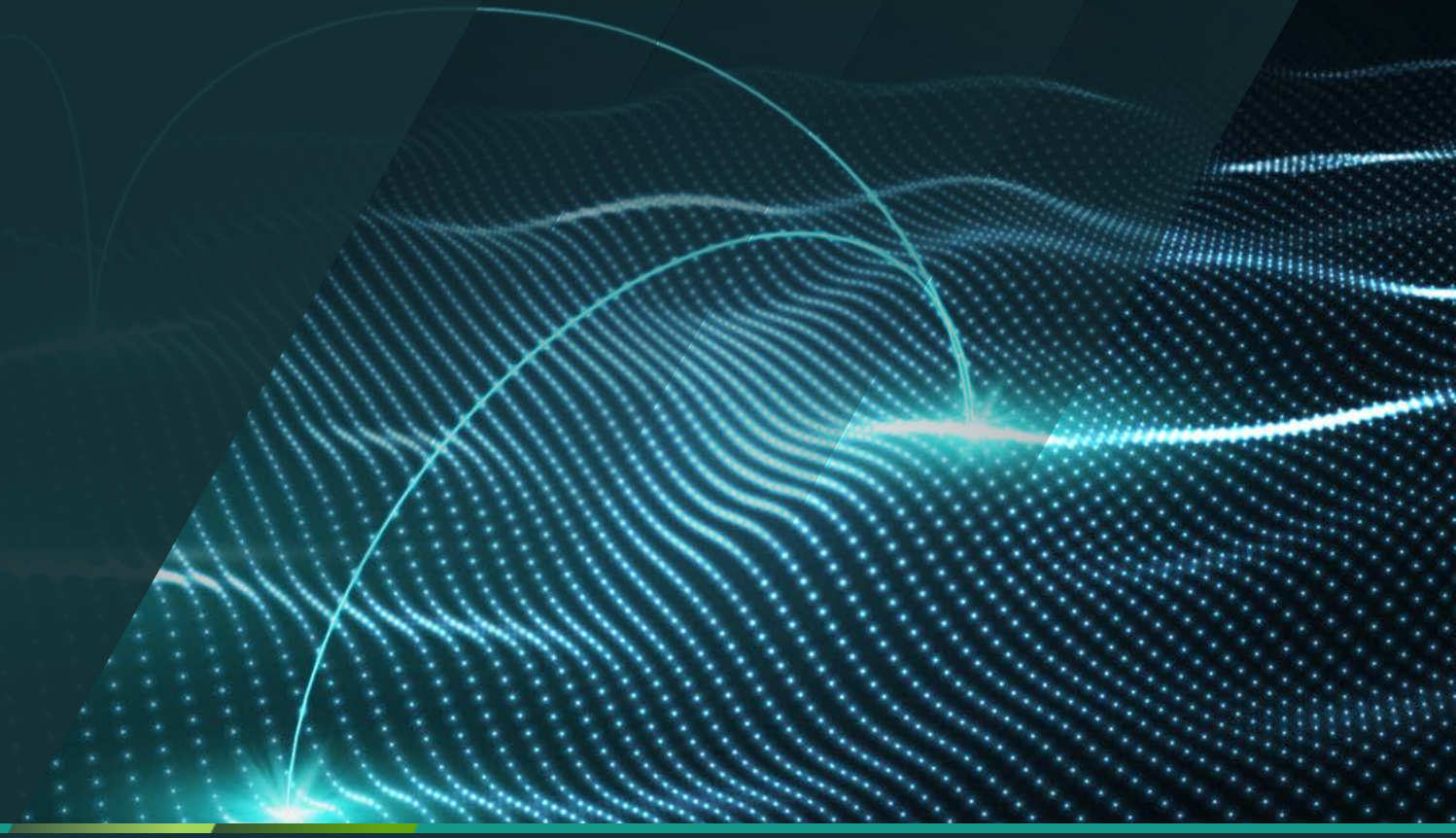
- Speculative Reads
- RCiEP vs DVSEC mapping
- Global Persistent Flush (GPF) + PMEM Support
- Power Budgeting
- Enhanced Error Handling + Isolation
  - Security violations
  - Surprise link down
- Skid Mode
- Link QoS

Wrap-up

- CXL represents a major change in server architecture.
- CXL 1.1 enables device-level memory expansion and coherent acceleration modes.
- CXL 2.0 augments CXL 1.1 with enhanced fanout support and a variety of additional features (some of which were reviewed in this webinar).
- CXL supporting platforms are due later this year.
- With CXL enabled applications around the corner, novel mechanisms efficiently utilizing vast scales of resources will enable a whole new world of compute.

Q&A

Please share your questions in the Question Box



Thank You