Introducing the CXL™ 3.0 Specification

Ishwar Agarwal, CXL Consortium

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

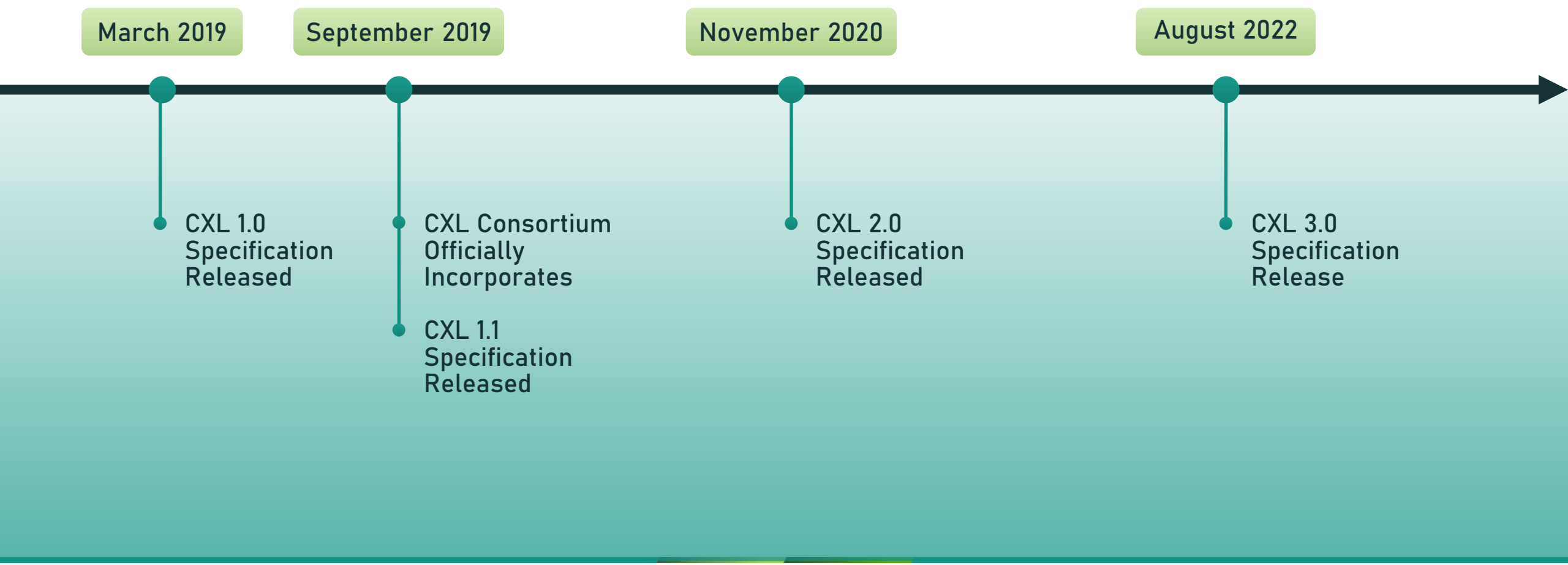
CXL Board of Directors

CXL Compute Express Link™

Industry Open Standard for High Speed Communications

200+ Member Companies

CXL Specification Release Timeline



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

CXL 3.0 Spec Feature Summary

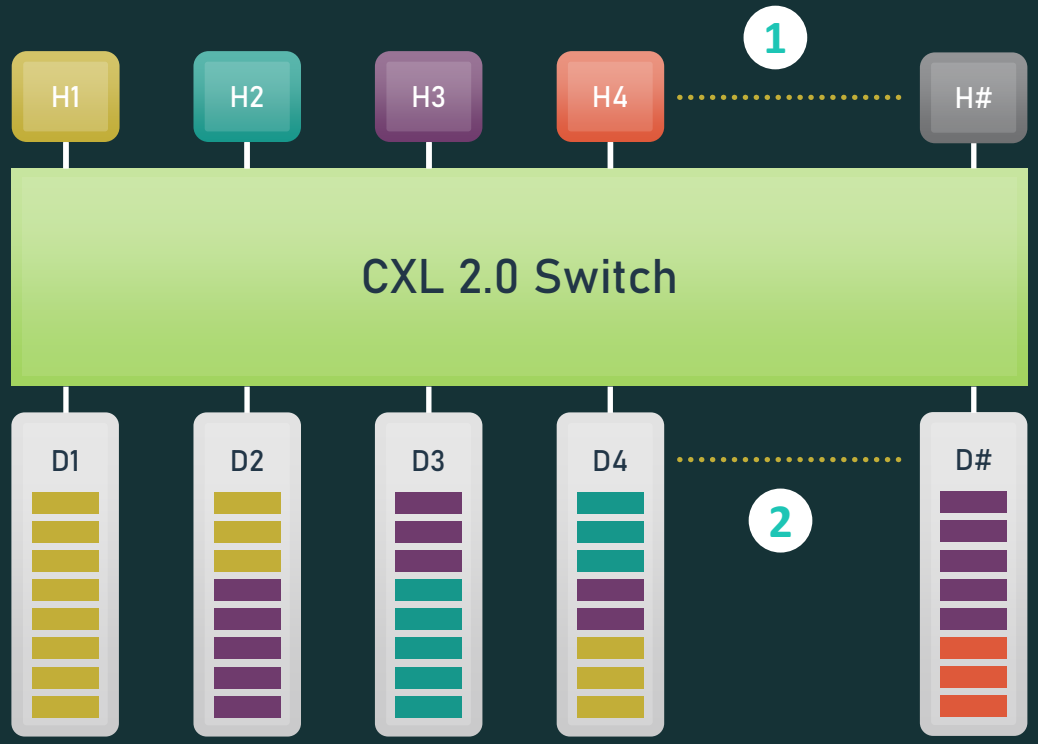
Features	CXL 1.0 / 1.1	CXL 2.0	CXL 3.0
Release date	2019	2020	1H 2022
Max link rate	32GTs	32GTs	64GTs
Flit 68 byte (up to 32 GTs)	✓	✓	✓
Flit 256 byte (up to 64 GTs)			✓
Type 1, Type 2 and Type 3 Devices	✓	✓	✓
Memory Pooling w/ MLDs		✓	✓
Global Persistent Flush		✓	✓
CXL IDE		✓	✓
Switching (Single-level)		✓	✓
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

RECAP: CXL 2.0 FEATURE SUMMARY

MEMORY POOLING

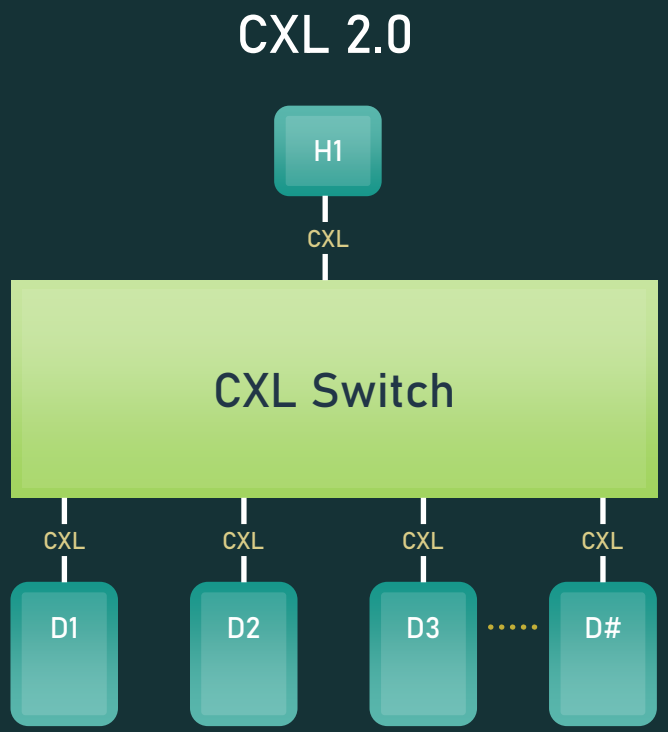


1. Device memory can be allocated across multiple hosts.

2. Multi Logical Devices allow for finer grain memory allocation

RECAP: CXL 2.0 FEATURE SUMMARY

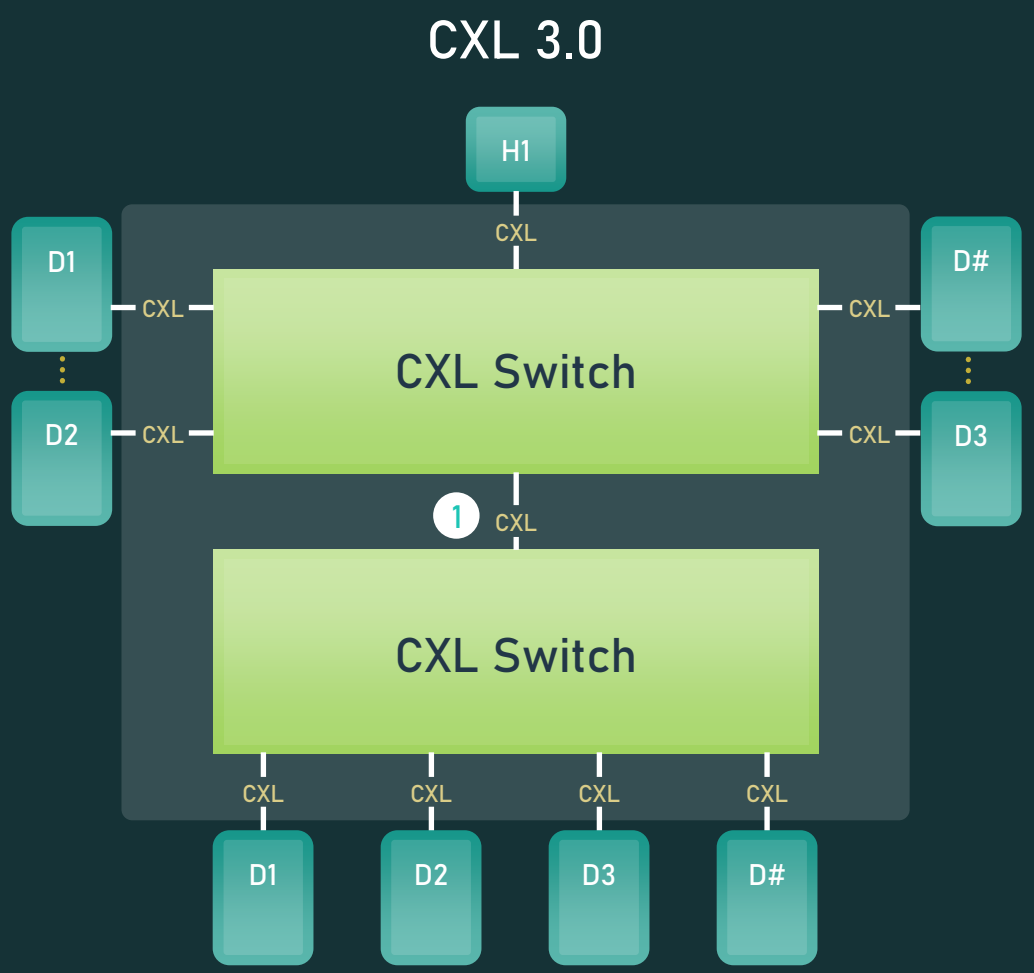
SWITCH CAPABILITY

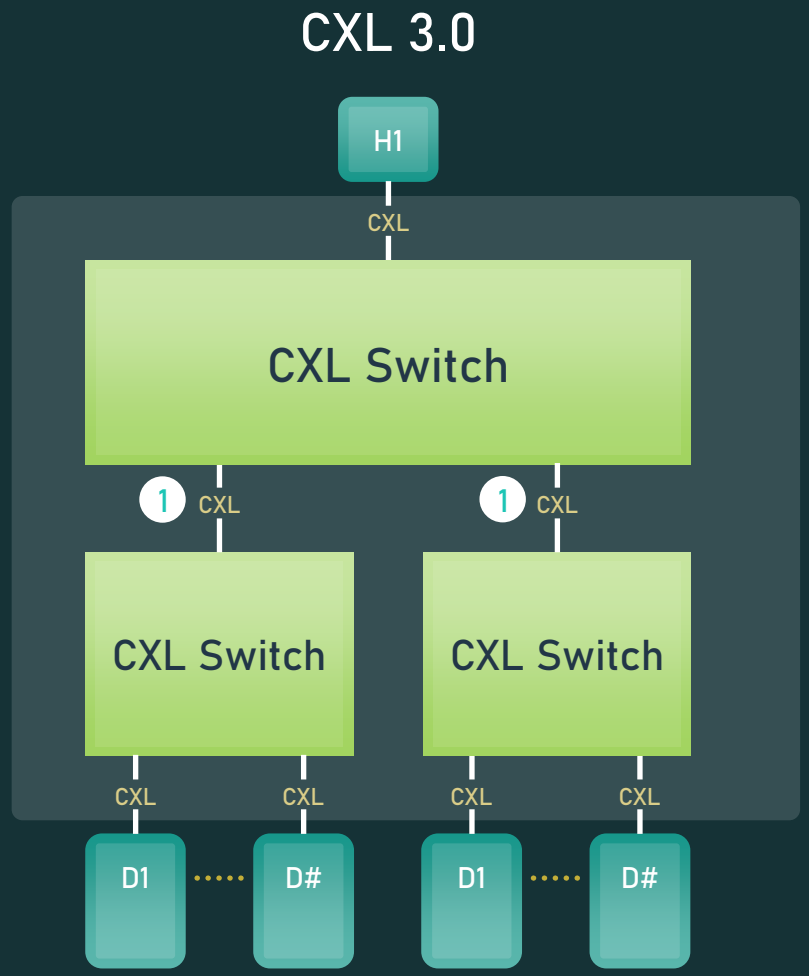


- Supports single-level switching
- Enables memory expansion and resource allocation

CXL 3.0: SWITCH CASCADE/FANOUT

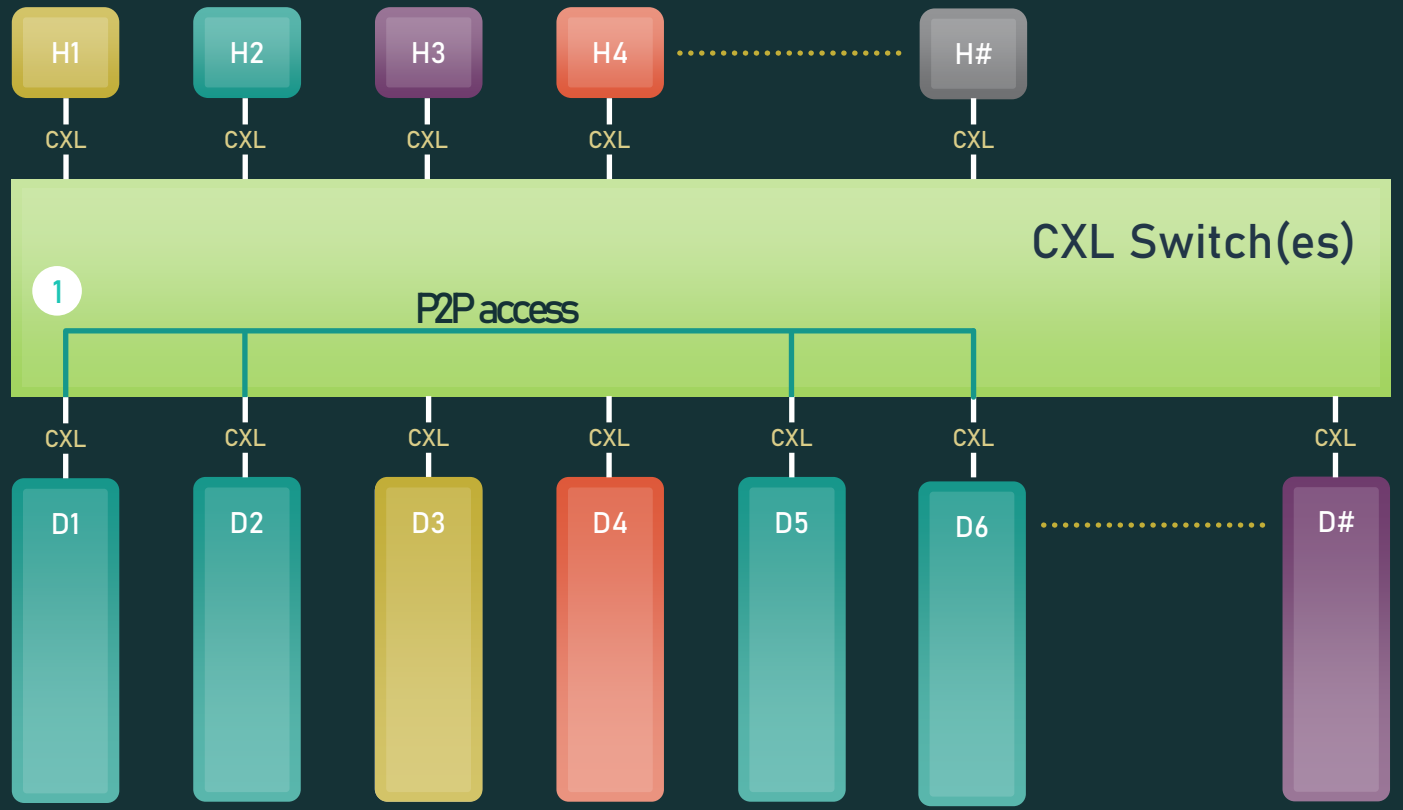
Supporting vast array of switch topologies





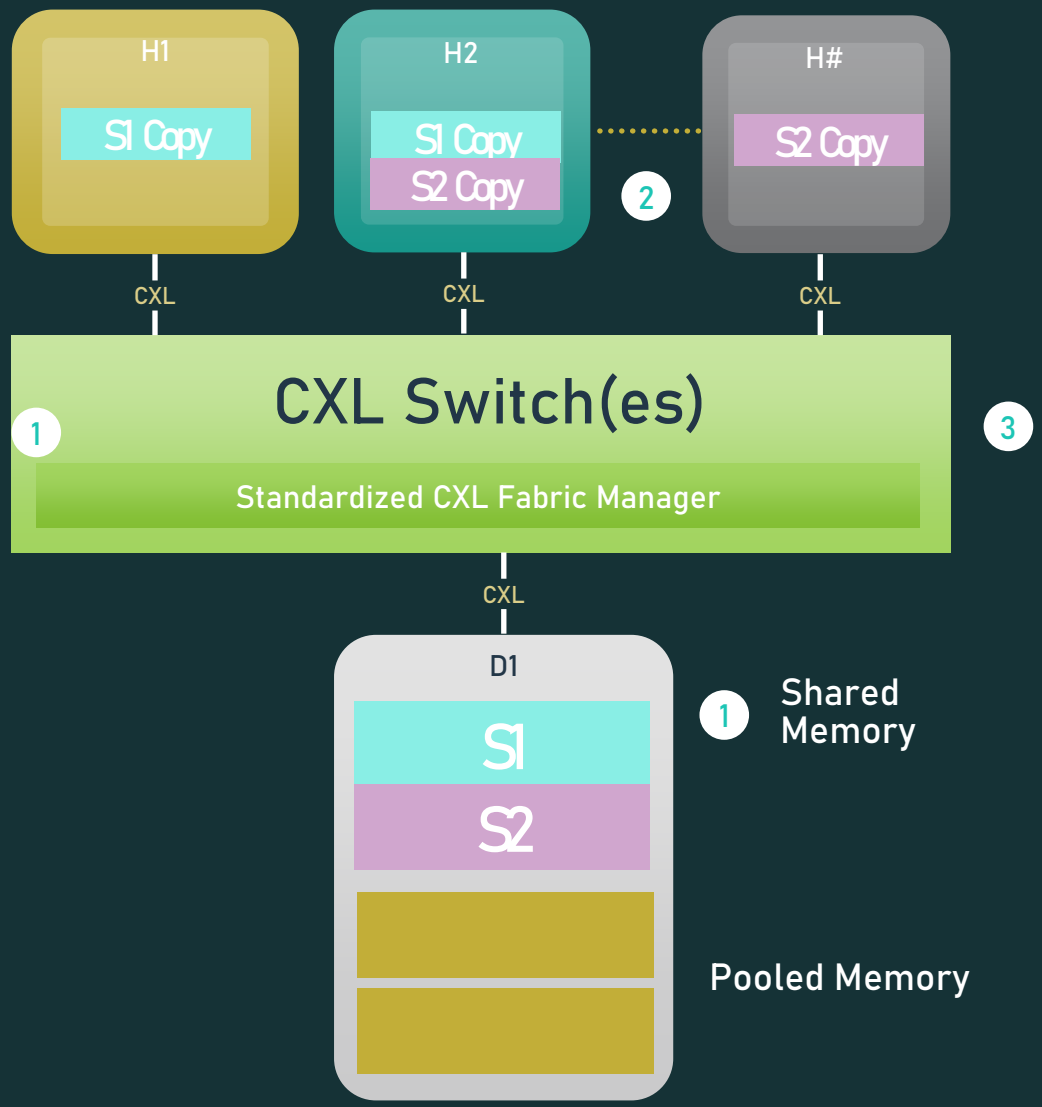
1 Multiple switch levels (aka cascade)
- Supports fanout of all device types

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

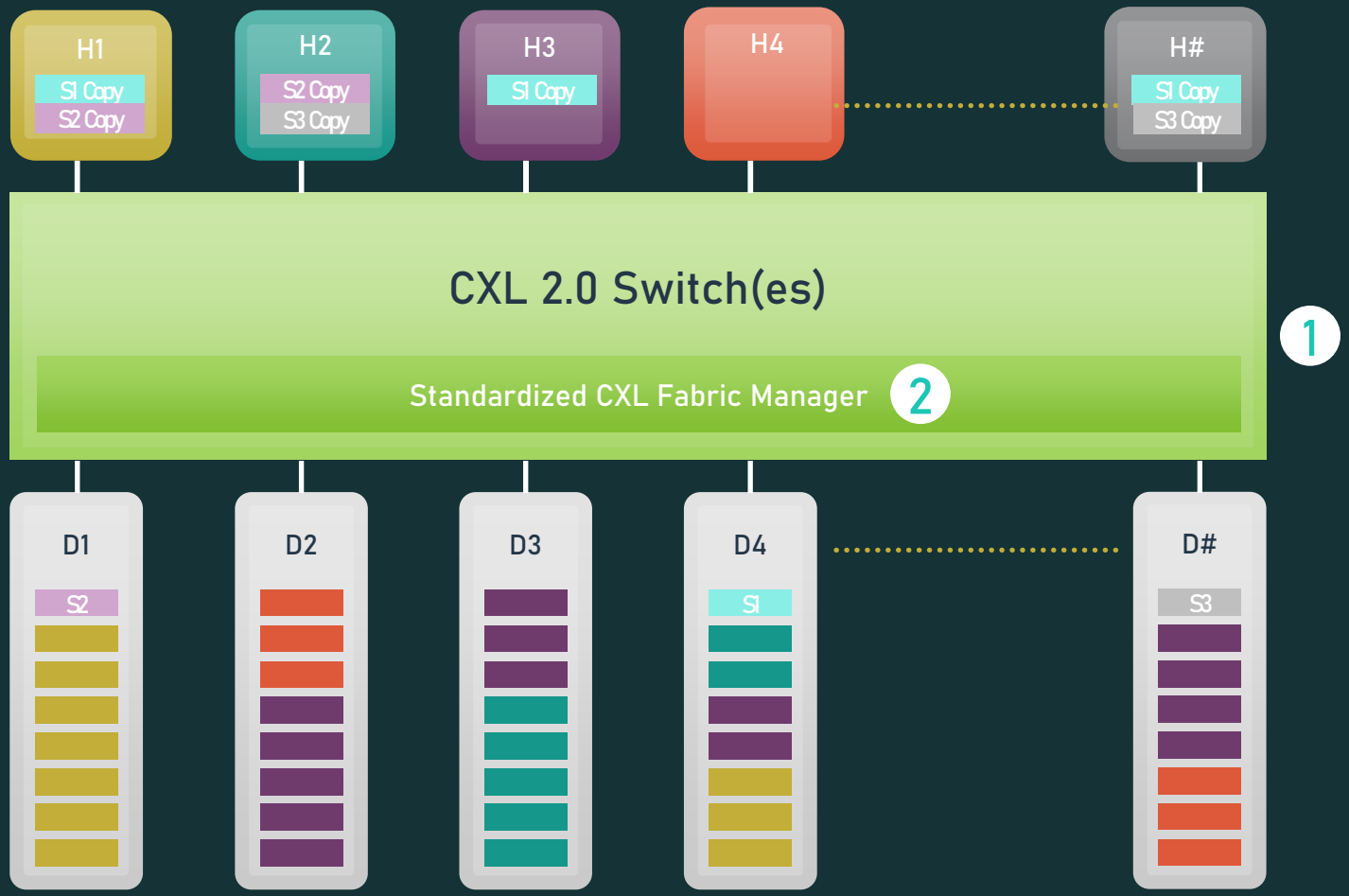


1. Device memory can be shared by all hosts to increase data flow efficiency and improve memory utilization

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

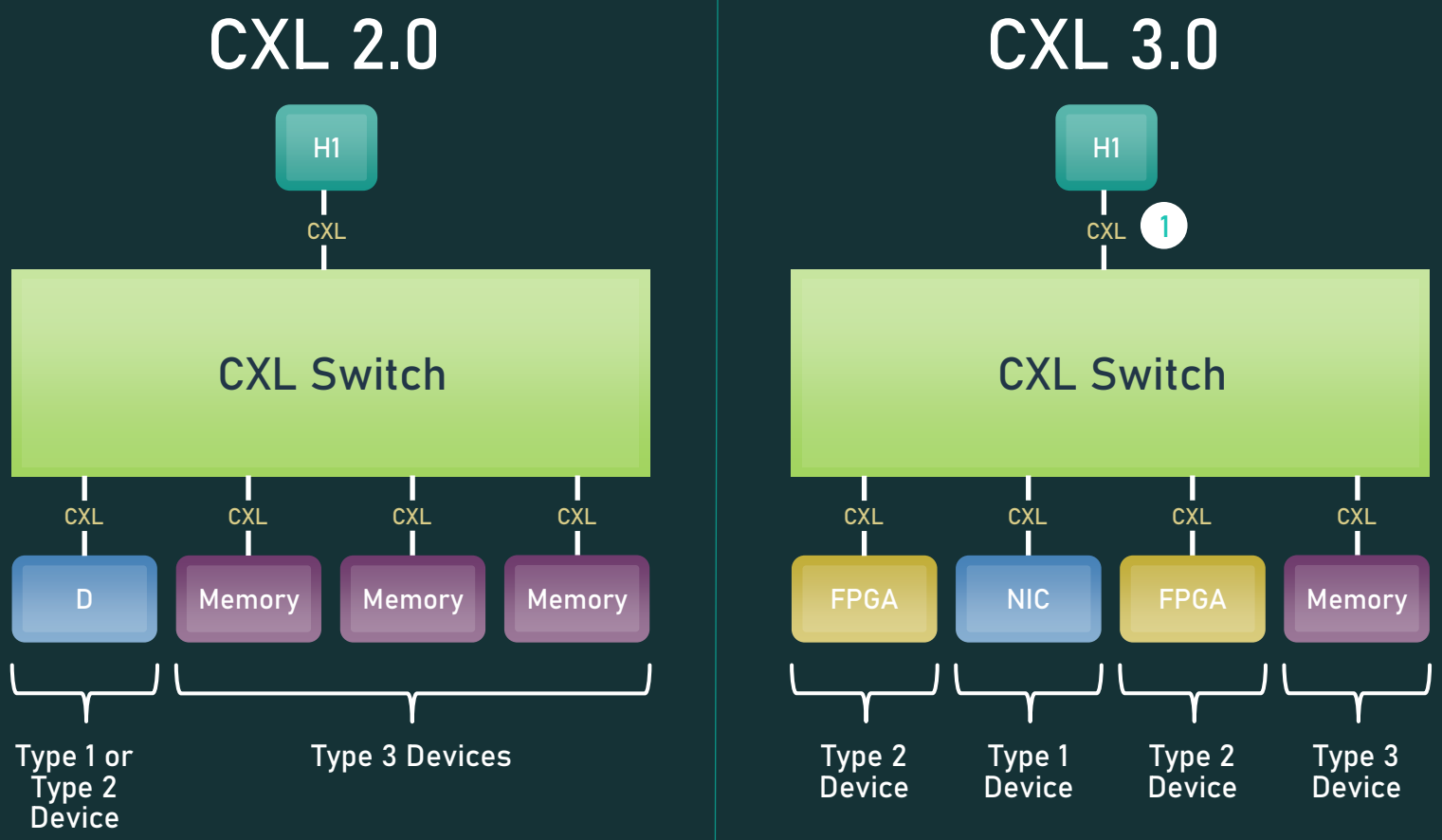
3. CXL 3.0 defined mechanisms to enforce hardware cache coherency between copies

CXL 3.0: POOLING & SHARING



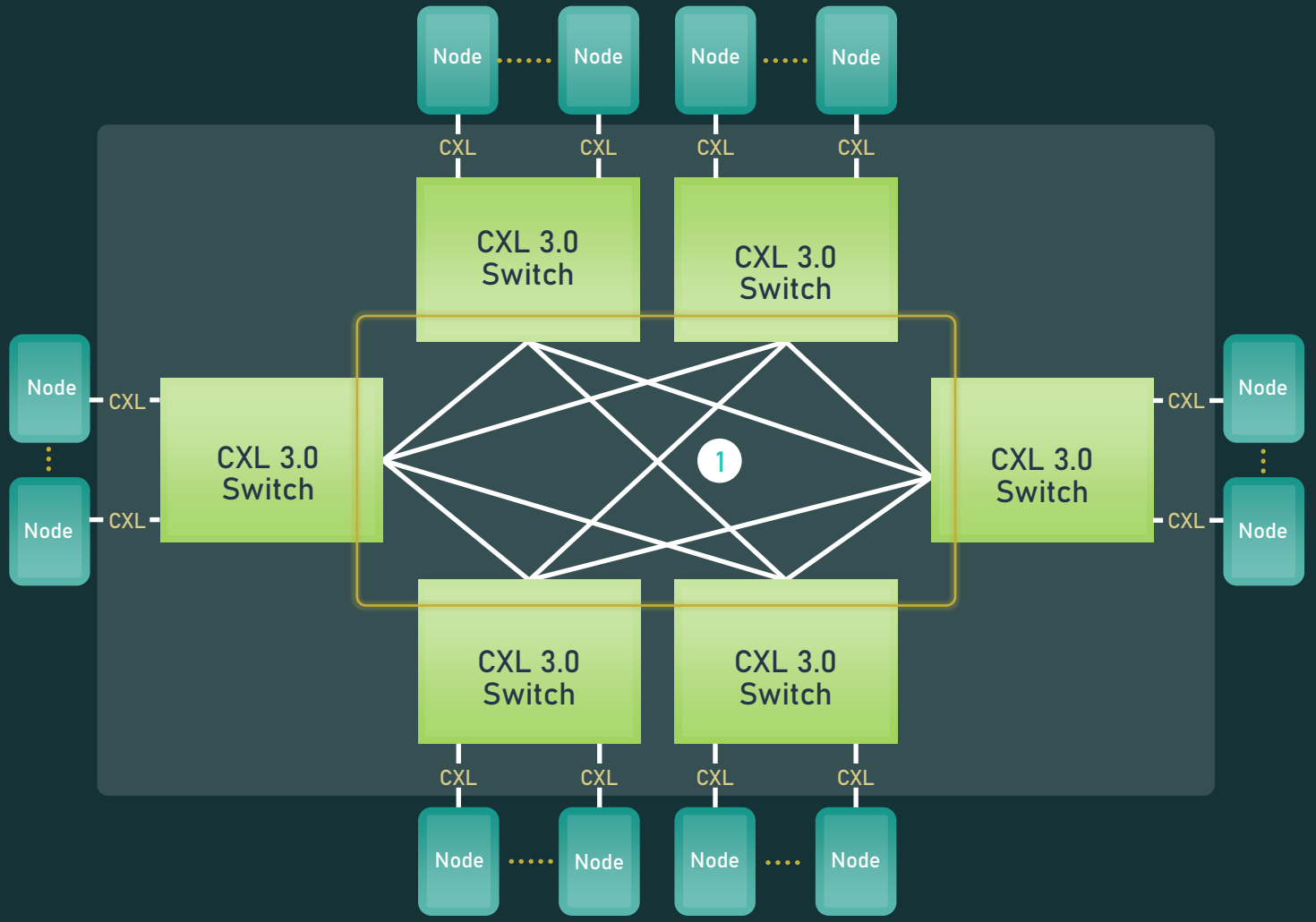
1. Expanded use case showing memory sharing and pooling
2. CXL Fabric Manager is available to setup, deploy, and modify the environment

CXL 3.0: MULTIPLE DEVICES OF ALL TYPES PER ROOT PORT



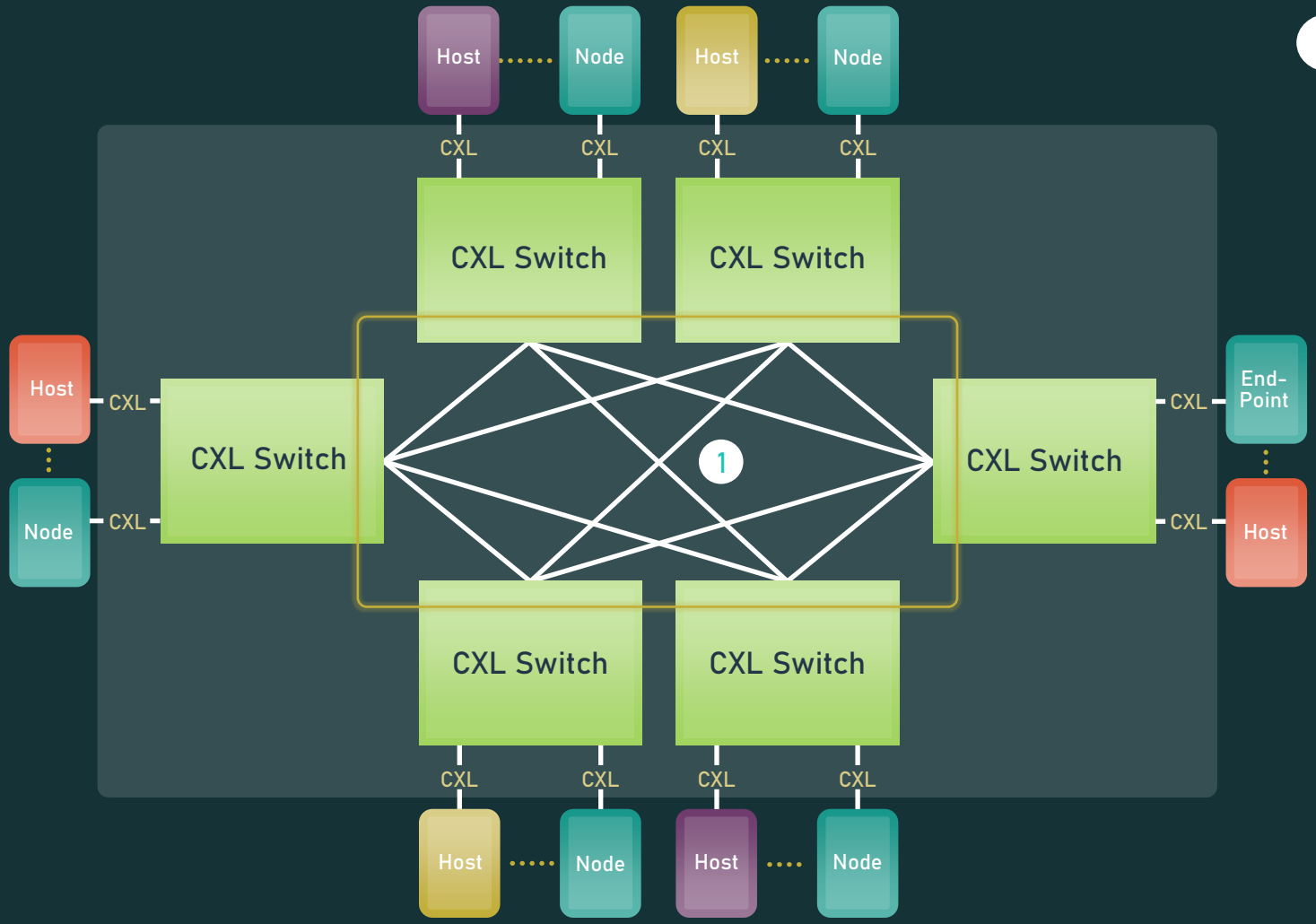
1 Each host's root port can connect to more than one device type

CXL 3.0: FABRICS OVERVIEW



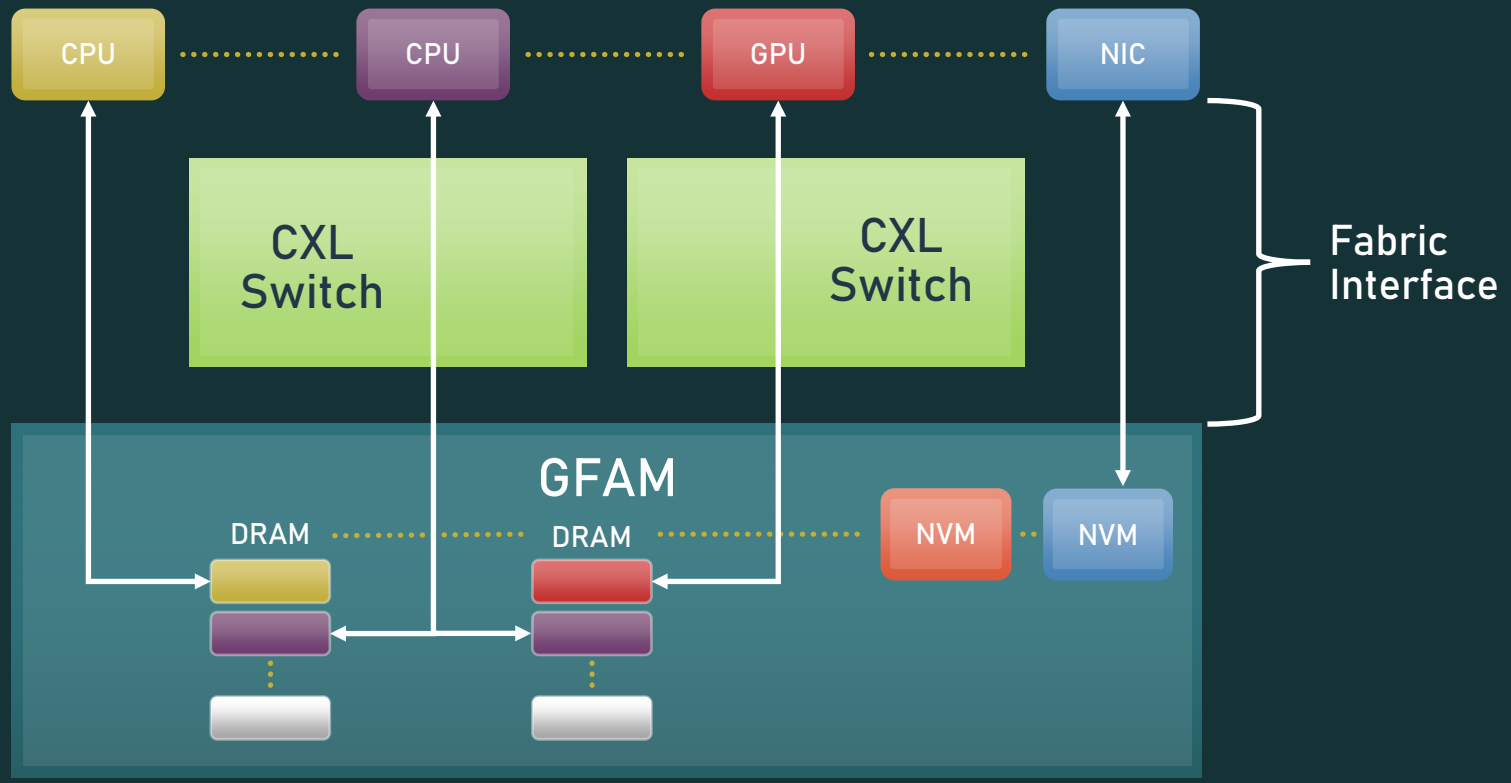
1. CXL 3.0 enables non-tree architectures
   - Each node can be a CXL Host, CXL device or PCIe device

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

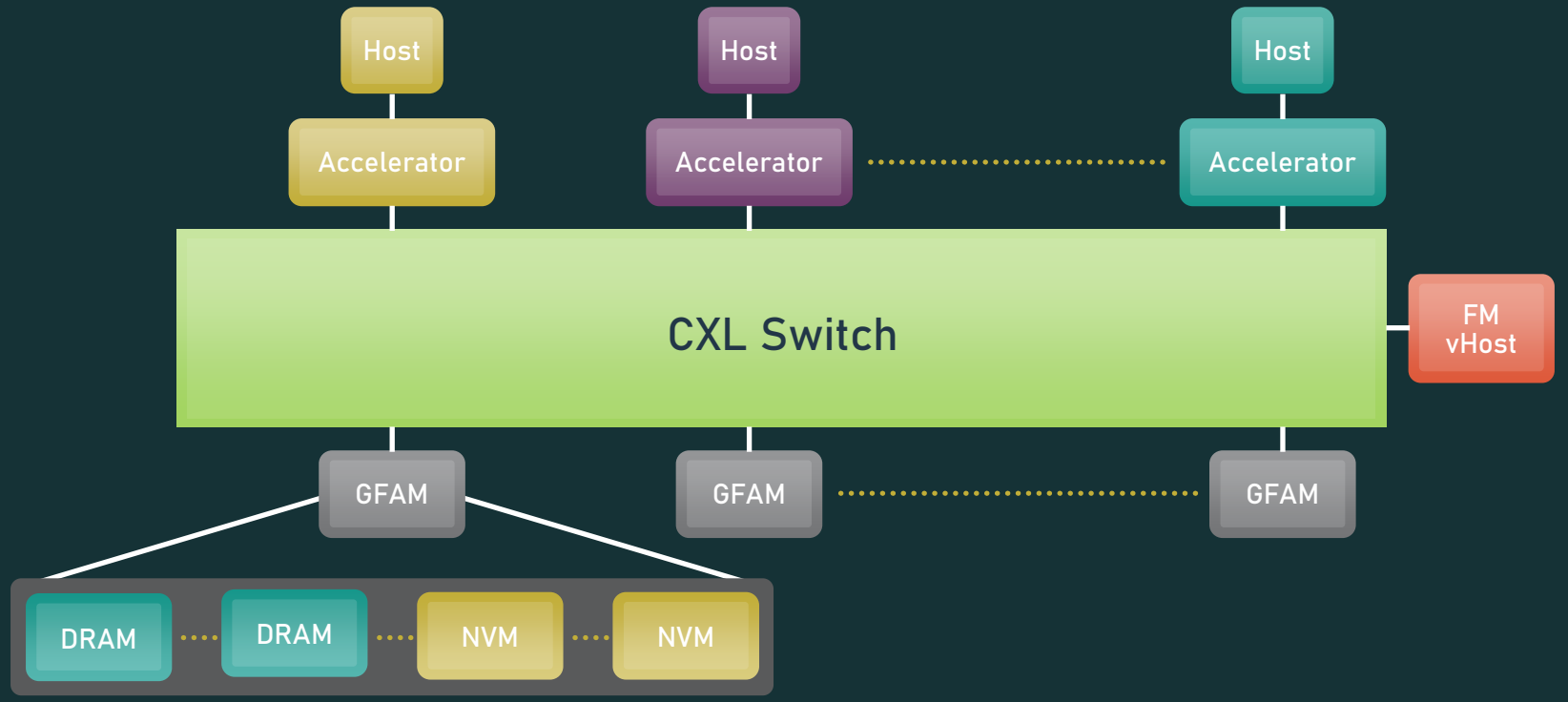
CXL 3.0: GLOBAL FABRIC ATTACHED MEMORY (GFAM) DEVICE



- CXL 3.0 enables Global Fabric Attached Memory (GFAM) architecture which differs from traditional processor centric architecture by disaggregating the memory from the processing unit and implements a shared large memory pool
- Memory can be of the same type or different types which can be accessed by multiple processors directly connected to GFAM or through a CXL switch

CXL 3.0: FABRICS EXAMPLE USE CASE

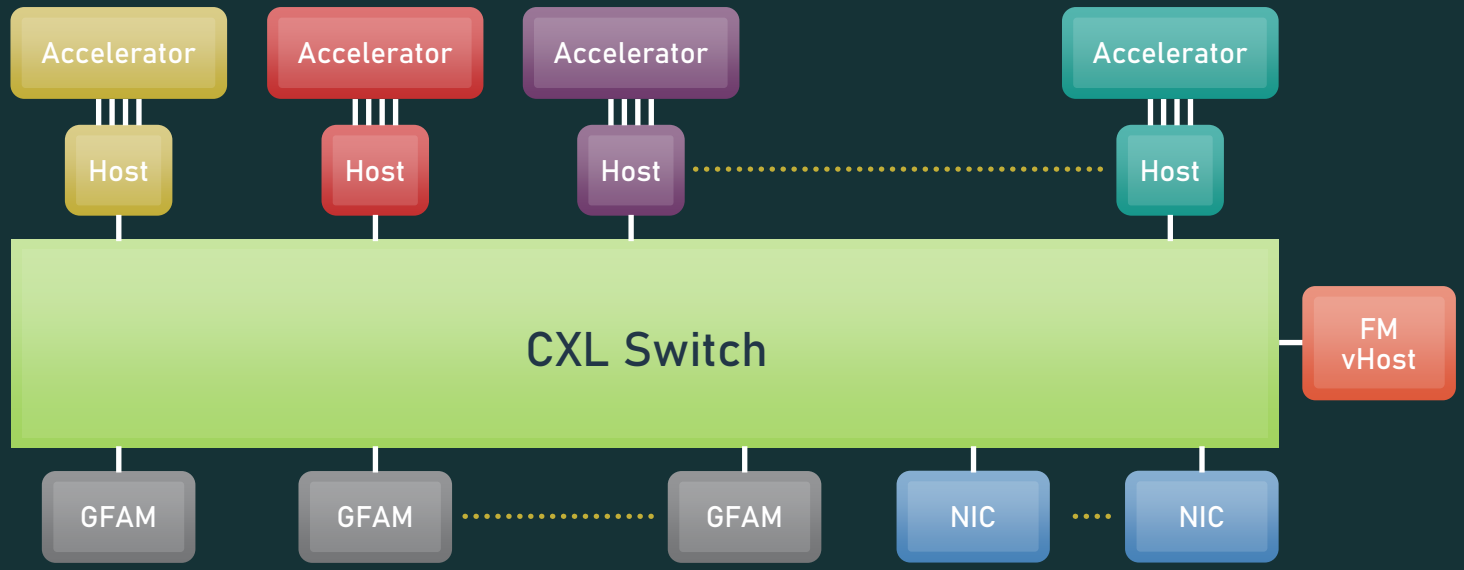
Machine Learning Accelerator and GFAM Device in a Fabric Architecture



GFAM enables multiple media types, i.e. DRAM, Flash, future memory types

CXL 3.0: FABRICS EXAMPLE USE CASE

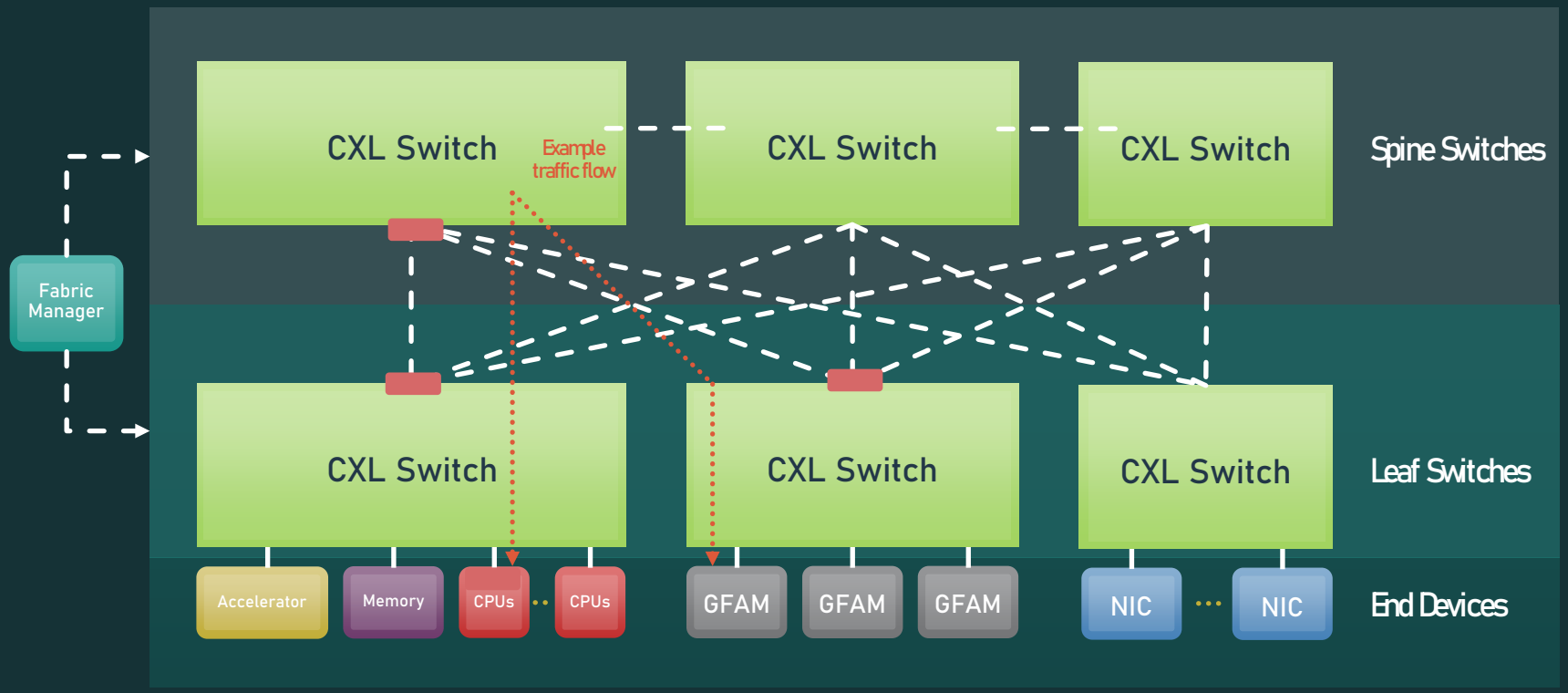
HPC/Analytics



Sharing memory and networking devices to reduce cost and improve efficiency

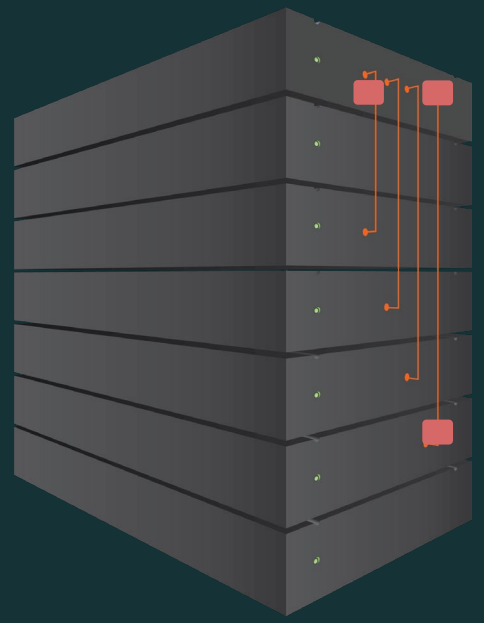
CXL 3.0: FABRICS EXAMPLE USE CASE

Composable Systems with Spine/Leaf Architecture



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



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 and visit ComputeExpressLink.org for additional resources
  - Follow us on Twitter and LinkedIn for updates!

Thank You