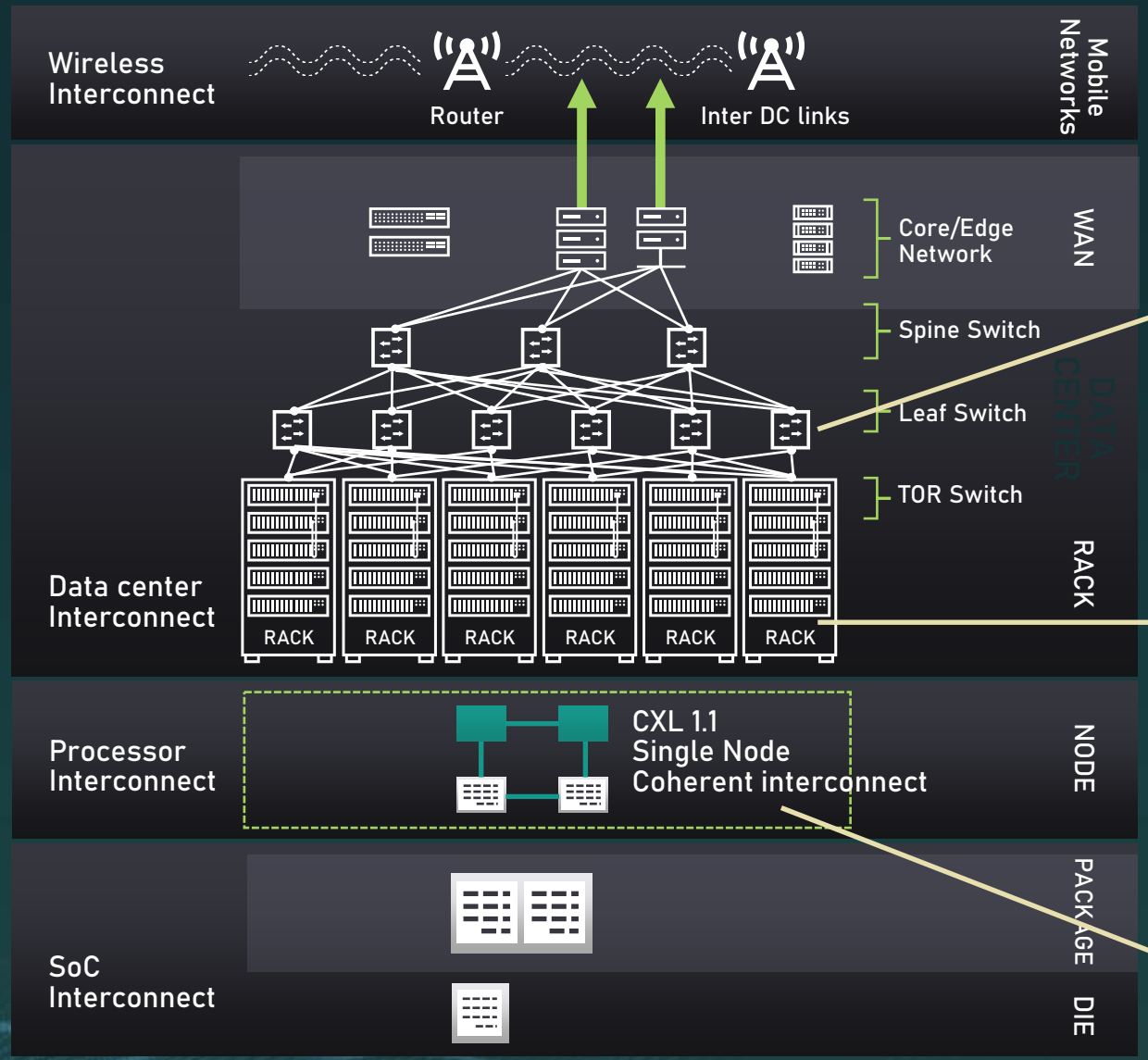
2024 OFA Virtual Workshop

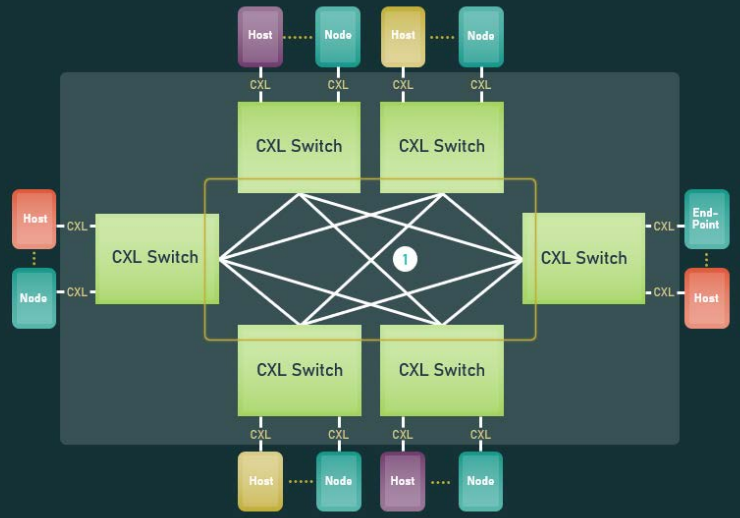
System Composability Using CXL

Kurtis Bowman, Marketing Working Group Co-Chair, CXL Consortium
Mitch Wright, CXL System Architect, Liqid

Data Center: Expanding Scope of CXL



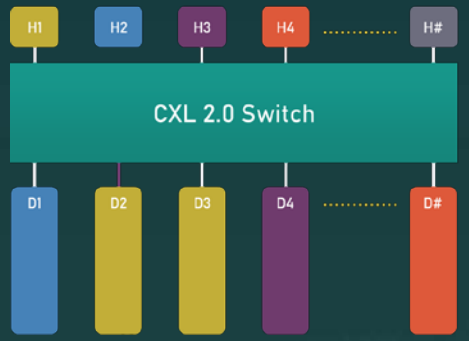
CXL 3.0/3.1



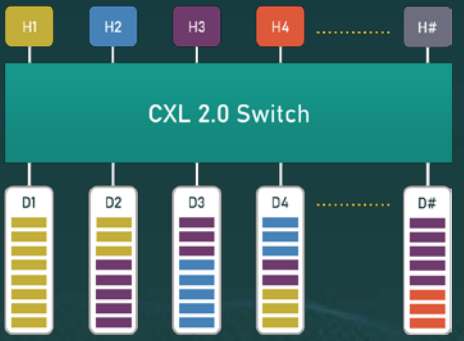
Composable Fabric growth for disaggregation/pooling/accelerator

CXL 2.0

Memory/Accelerator Pooling with Single Logical Devices



Memory Pooling with Multiple Logical Devices



Multiple Nodes inside a Rack/ Chassis supporting pooling of resources

CXL Specification Feature Summary

Features	CXL 1.0 / 1.1	CXL 2.0	CXL 3.0	CXL 3.1
Release date	2019	2020	August 2022	November 2023
Max link rate	32GTs	32GTs	64GTs	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)			✓	✓
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 34-bit of meta data, RAS capability enhancements)				✓

Not Supported ✓ Supported

CXL 3.1 Fabric Management Features

CXL 3.1 Fabric Management Features

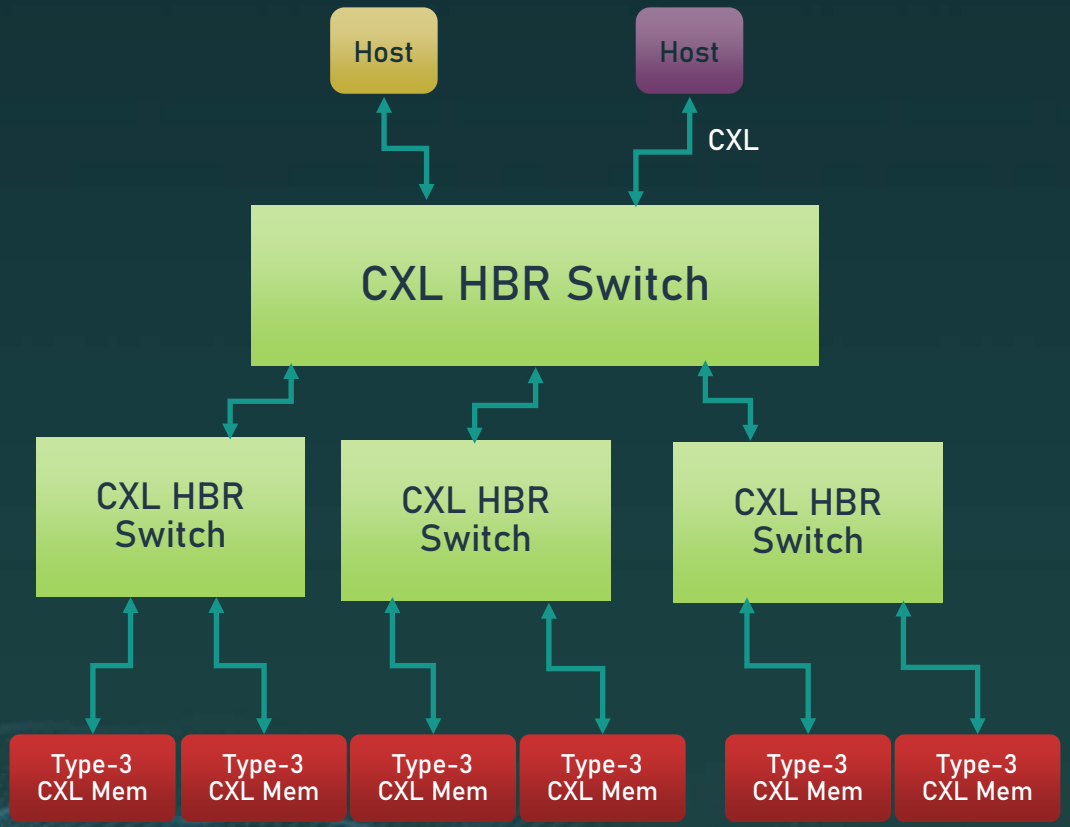
- Fabric Decode/Routing requirements
- Host-to-Host communication with Global Integrated Memory (GIM) concept
- Direct P2P .mem support through PBR Switches
  - Adds symmetric Link Layer definition
  - Enables direct caching of CXL.mem for an accelerator, which is not possible with Unordered IO (UIO)
- Fabric Manager API definition for PBR Switch

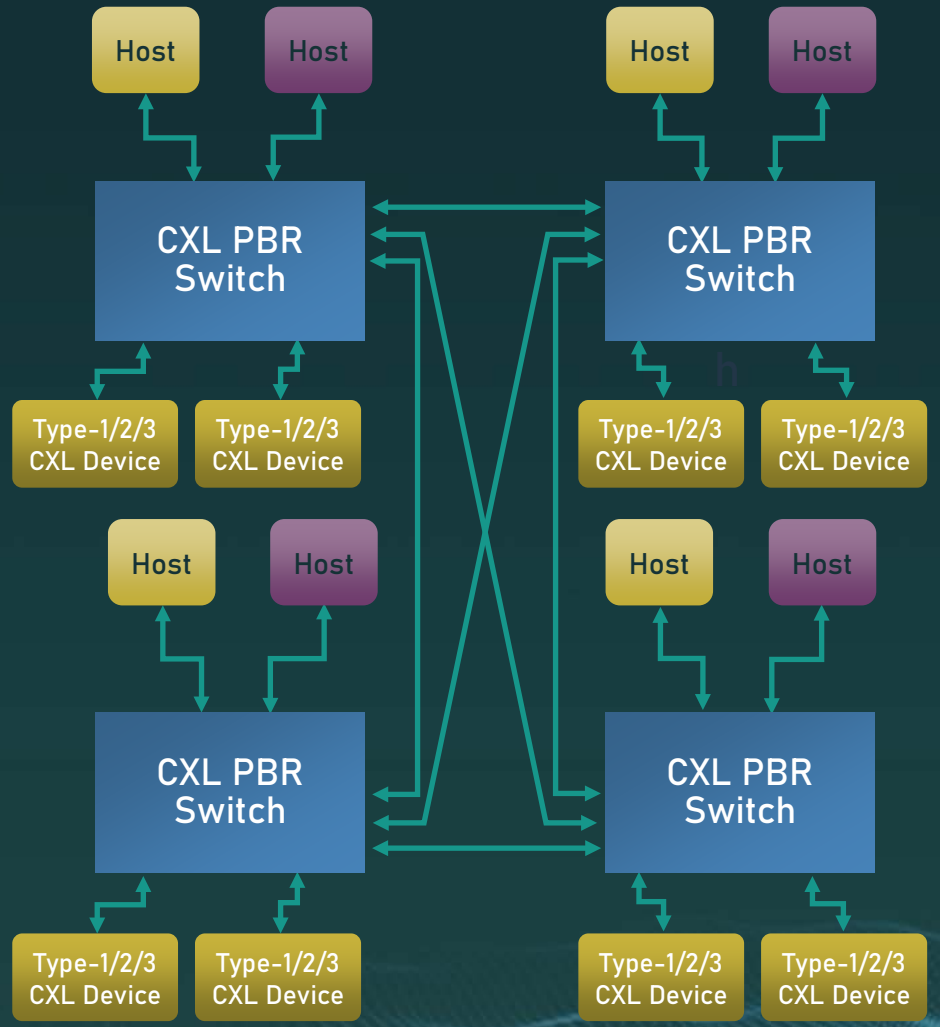
CXL 3.1: Fabric Management Features

Port-Based Routing (PBR) compared to Hierarchy-Based Routing (HBR)

Support fabric topologies other than tree topologies that HBR switches offer

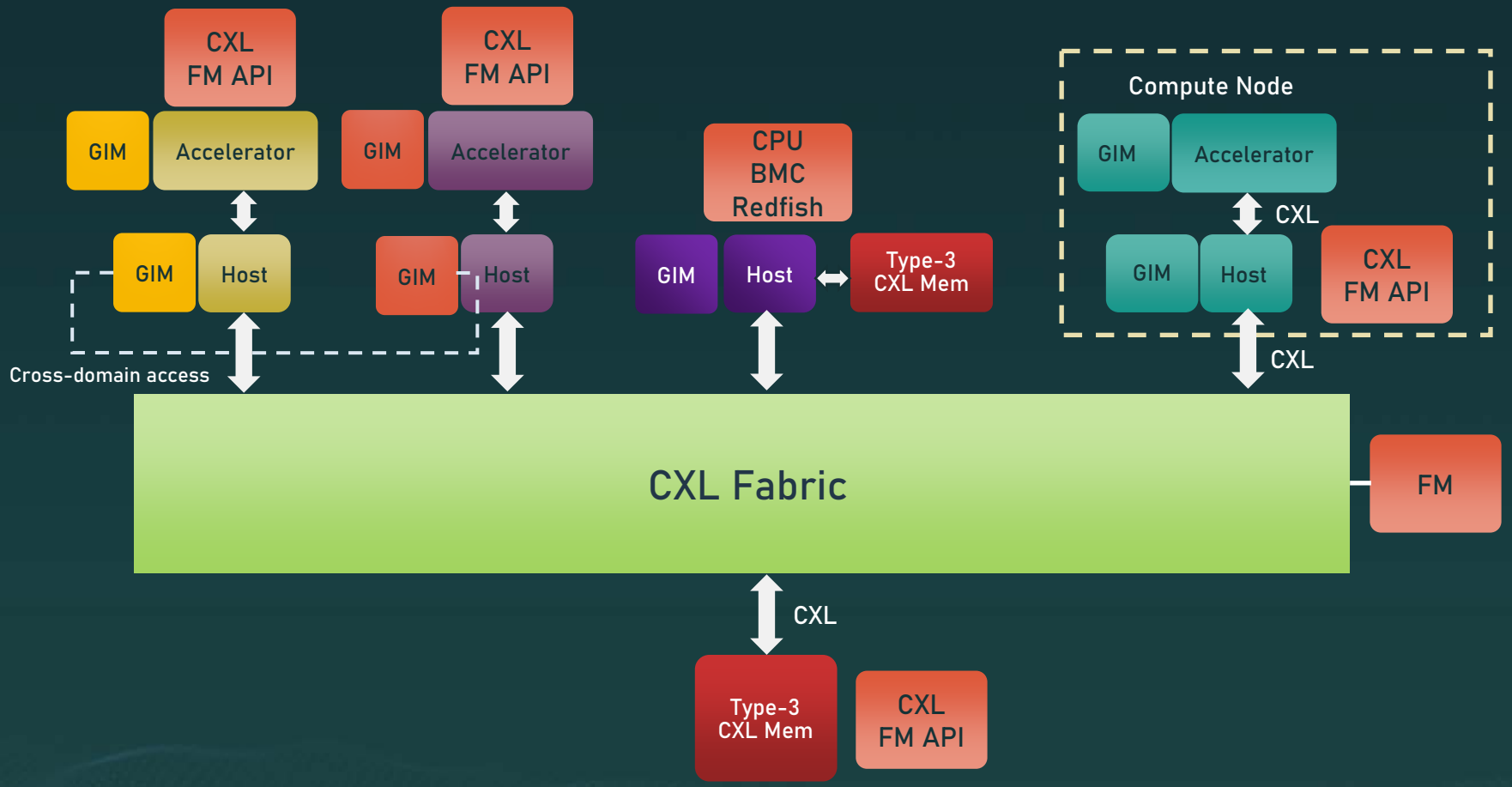
- Address-based, non-prescriptive routing for large memory fabrics
- Supports tree, mesh, ring, star, butterfly, and multi-dimensional topologies





CXL 3.1: Fabric Management Features

Host-to-Host communication with Global Integrated Memory (GIM) concept

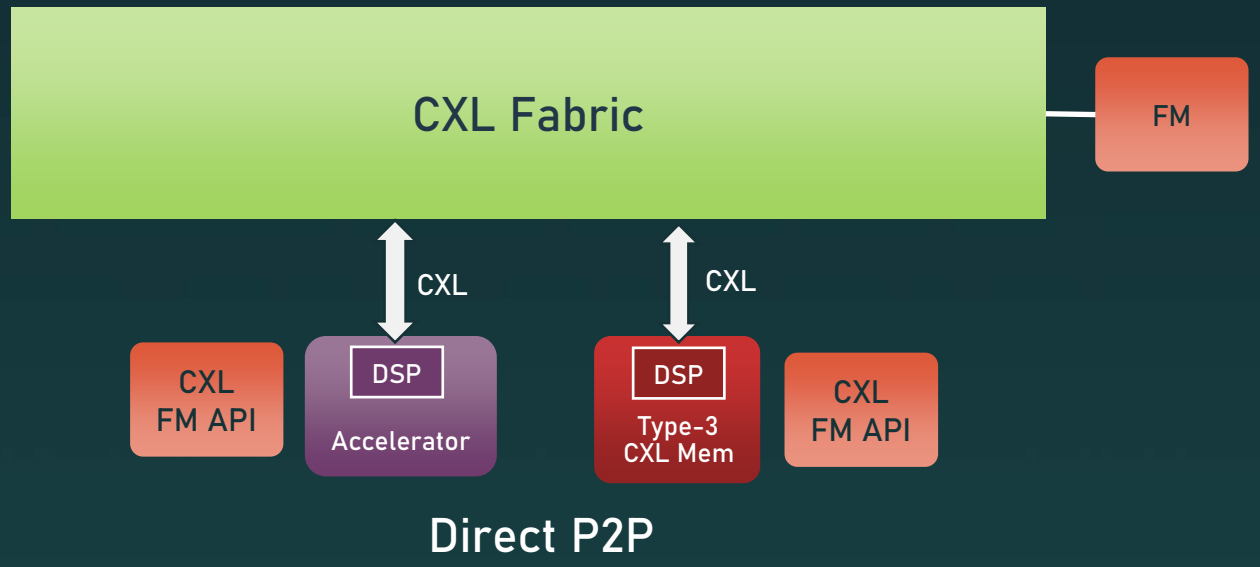


Host to Fabric-Attached Memory (FAM) communication with Global Integrated Memory (GIM)

- Multiple Hosts mapping CXL Fabric-Attached Memory devices
- Hosts and FAM devices can initiate cross-domain accesses to GIM

CXL 3.1: Fabric Management Features

Direct peer-to-peer (P2P) .mem support for Accelerators through PBR Switches



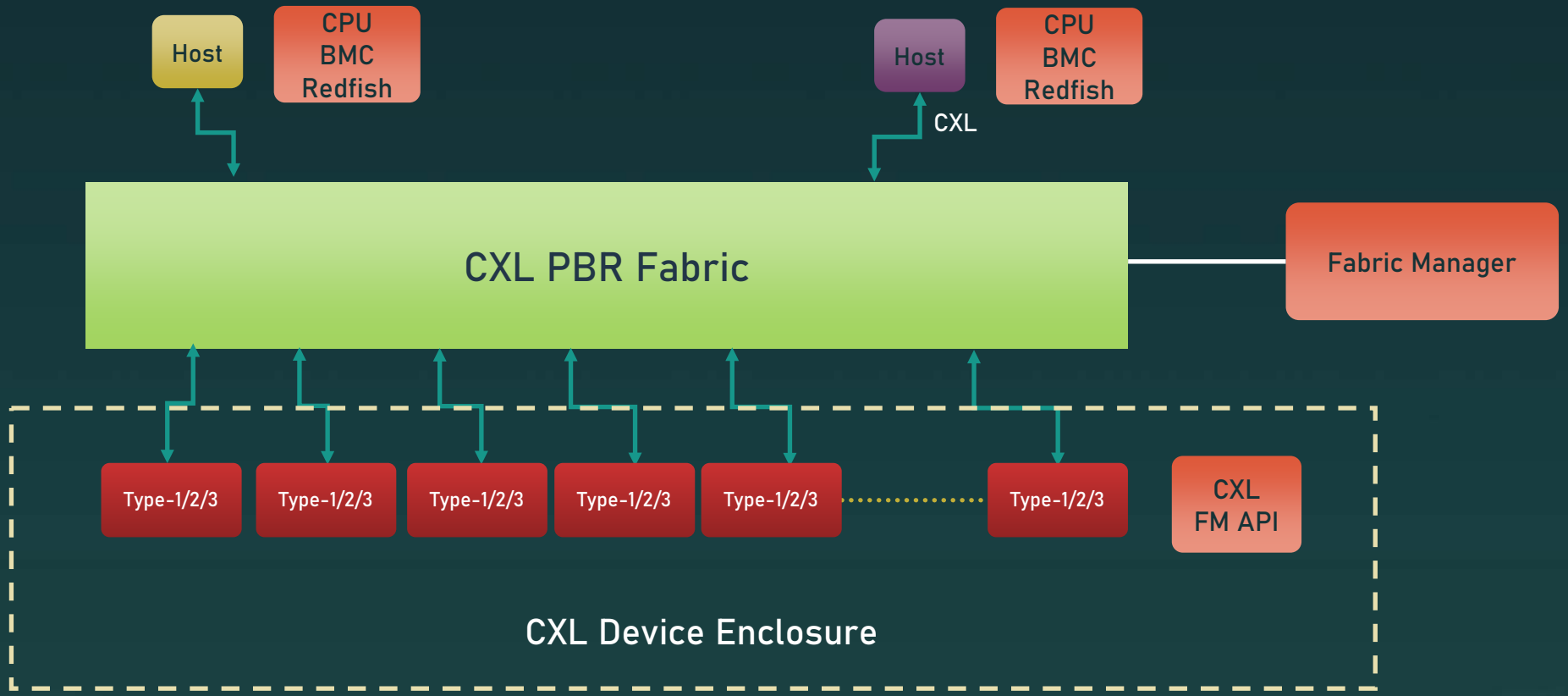
Direct P2P

Direct P2P CXL.mem for Accelerators

- Enables accelerator access to peer Type-3 memory
- The accelerator and Type-3 device must each be directly connected to an Edge Downstream Port (DSP)
- Utilizes port-based routing (PBR) for transactions

CXL 3.1: Fabric Management Features

Fabric Manager (FM) API definition for PBR Switch



CXL 3.1 Fabric Management Features

- CXL Fabric Management Integration with Industry Standards
  - DMTF Redfish®
    - Redfish Support for CXL
    - CXL to Redfish Mapping Specification
  - SNIA Swordfish®
    - SNIA Swordfish Specification
    - What is the SNIA Swordfish Standard
  - OFA Sunfish™
    - OFA Sunfish

Thank You

www.ComputeExpressLink.org