CXL 3.0: Novel Device Types, Capabilities, and Interconnects

EMPOWERING OPEN.

OCP GLOBAL SUMMIT

OCTOBER 18-20, 2022
SAN JOSE, CA

CXL 3.0: Novel Device Types, Capabilities, and Interconnects

Danny Moore, Senior Manager, Product Management and Strategy, Rambus

OPEN PLATINUM™

Agenda

- CXL 3.0 Overview
- Multi Headed Devices
- Dynamic Capacity Devices
- Fabrics

CXL 3.0 Overview

CXL 3.0 Specification

Fabric capabilities and management

Improved memory sharing and pooling

Symmetric coherency

Peer-to-peer

Expanded capabilities for increasing scale and optimizing resource utilization

- Fabric capabilities and fabric attached memory
- Enhance fabric management framework
- Memory pooling and sharing
- Peer-to-peer memory access
- Multi-level switching
- Near memory processing
- Multi-headed devices
- Multiple Type 1/Type 2 devices per root port
- Fully backward compatible to CXL 2.0, 1.1, and 1.0
- Supports PCIe® 6.0

CXL 3.0 Spec Feature Summary

Features	CXL 1.0 / 1.1	CXL 2.0	CXL 3.0
Release date	2019	2020	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			✓
Symmetric coherency (256 byte flit)			✓
Memory sharing (256 byte flit)			✓
Multiple Type 1/Type 2 devices per root port			✓
Fabric capabilities (256 byte flit)			✓

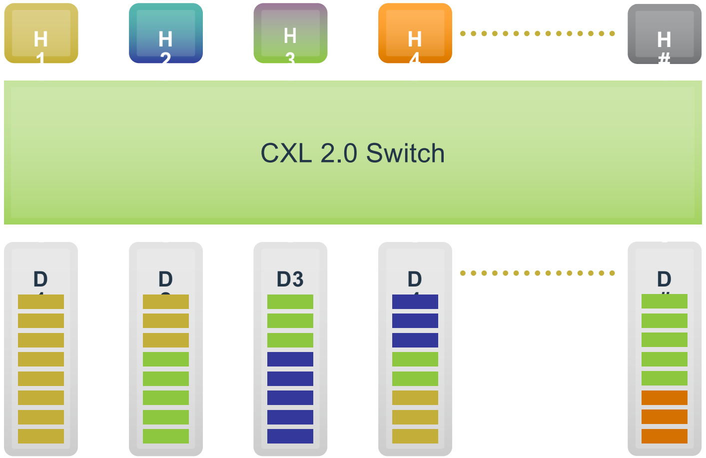
Not supported

✓ Supported

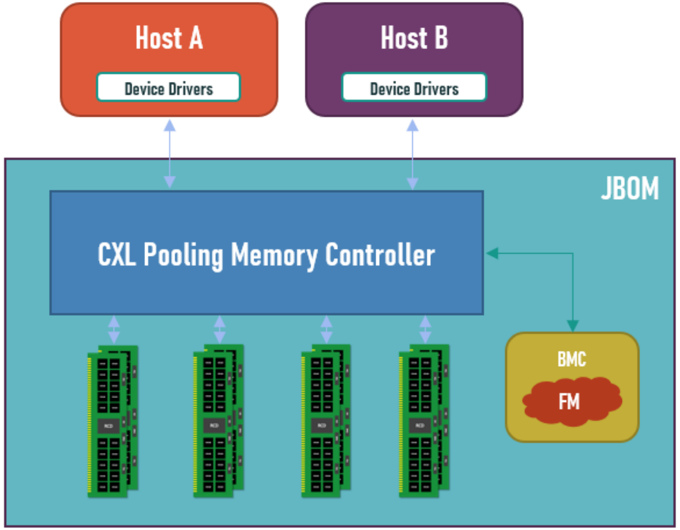
Multi-Headed Devices (MHDs)

Pooling in 2.0 & 3.0

2.0 – MLD and switch enabled pooling



3.0 – MHD

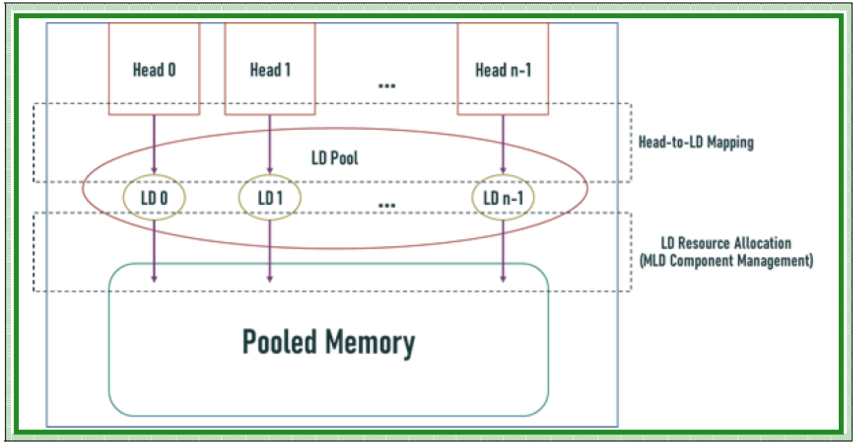


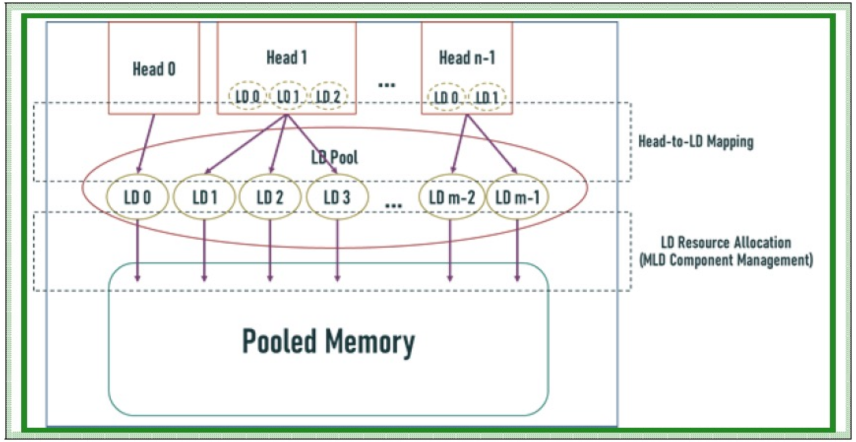
Multi Headed Devices (MHDs)

- **Two types of MHD**
  1. Multi-Headed-Single Logical Device (MH-SLD)
  2. Multi-Headed-Multiple Logical Device (MH-MLD)
- **MH-SLDs/MLDs are differentiated between how resources are presented to device heads**
- **MH-SLD presents an SLD to each head**
  - 1-1 mapping of LDs to heads
  - Heads can direct attach to different hosts
- **MH-MLD presents an MLD to each head**
  - Up to 16 LDs mapped to a single head
  - Requires a switch for MLD functionality

MH-SLD & MH-MLD

- Logical abstraction showing LD mapping to heads
  - MH-SLD showing a 1-1 mapping of a single LD to each head
  - MH-MLD showing additional composability with up to 16 LDs mapped to each head

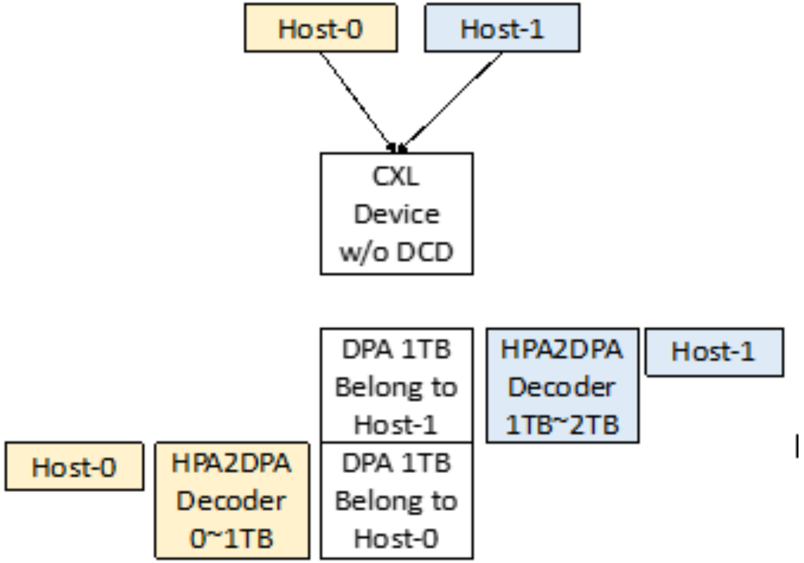




Dynamic Capacity Device (DCD)

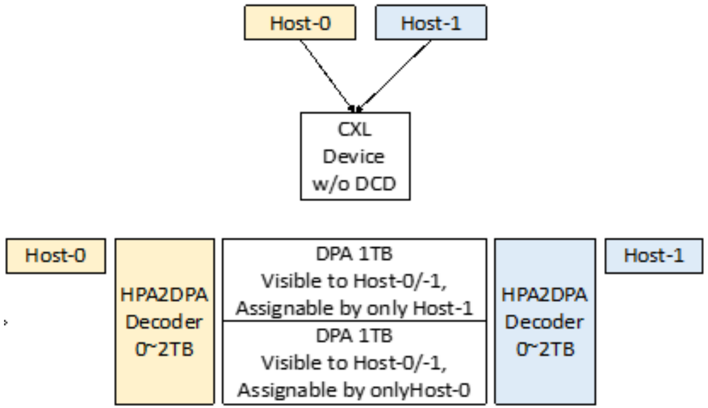
Problem Statement

- Prior to the definition of DCD, allocating and deallocating memory was very disruptive
- Capacity is added or released by reprogramming the HDM decoders
- For the new capacity to be utilized by the host, traffic must be quiesced and a system reset is likely needed to ensure that all memory is mapped accordingly



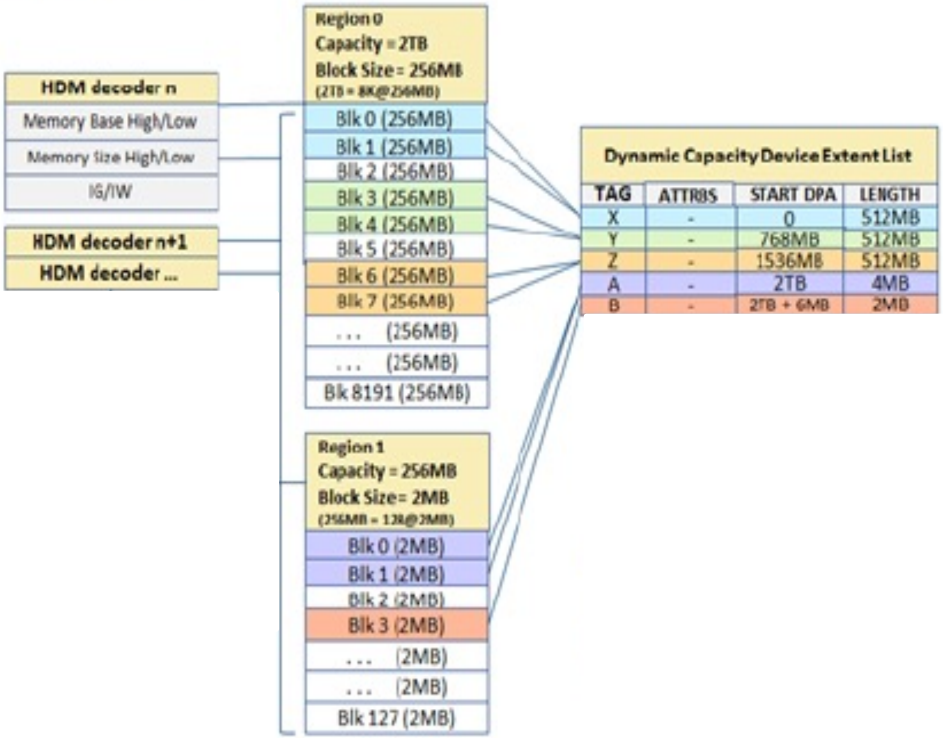
Dynamic Capacity Devices

- A Dynamic Capacity Device (DCD) is a memory device that implements the DCD command set
- Dynamic capacity is a feature that allows memory capacity to change dynamically without the need to reprogram the HDM decoders
- The DCD presents its maximum capacity to each host, HDM decoders are programmed for the entire DPA range, and the DCD command set is used by the host to discover the actual memory allocation



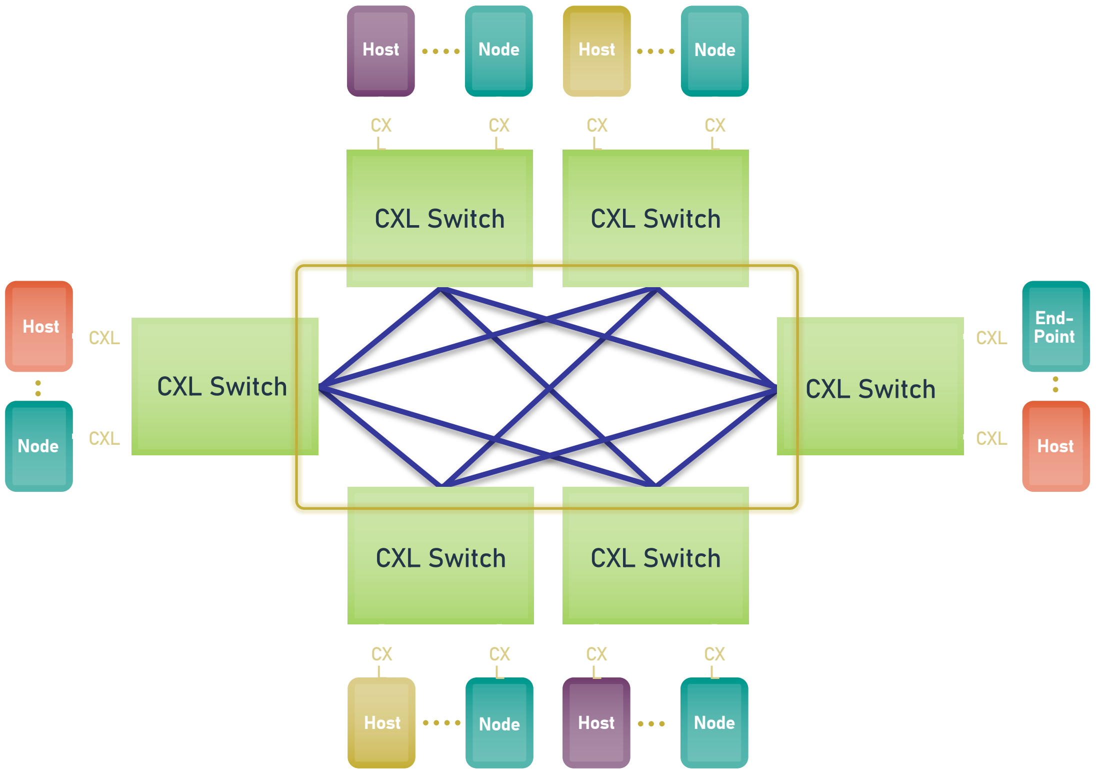
Dynamic Capacity Devices

- The Fabric Manager (FM) uses the DCD command set to query and configure the DCD
- The DPA is divided into 1-8 separate regions, and each region is subdivided by the DCD into fixed-size blocks
- The DCD communicates an extent list to each host describing all blocks that can be accessed



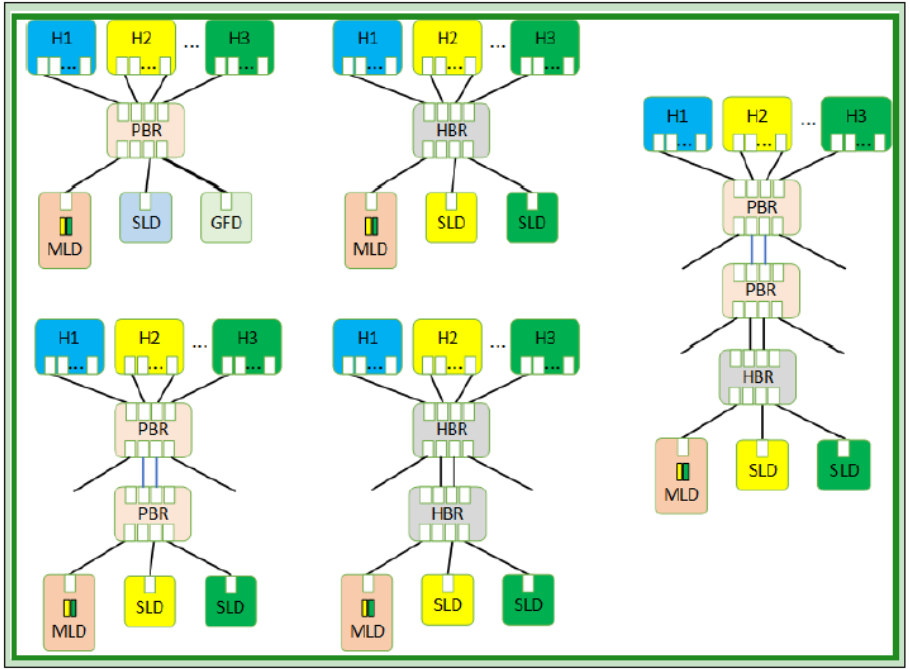
Fabrics

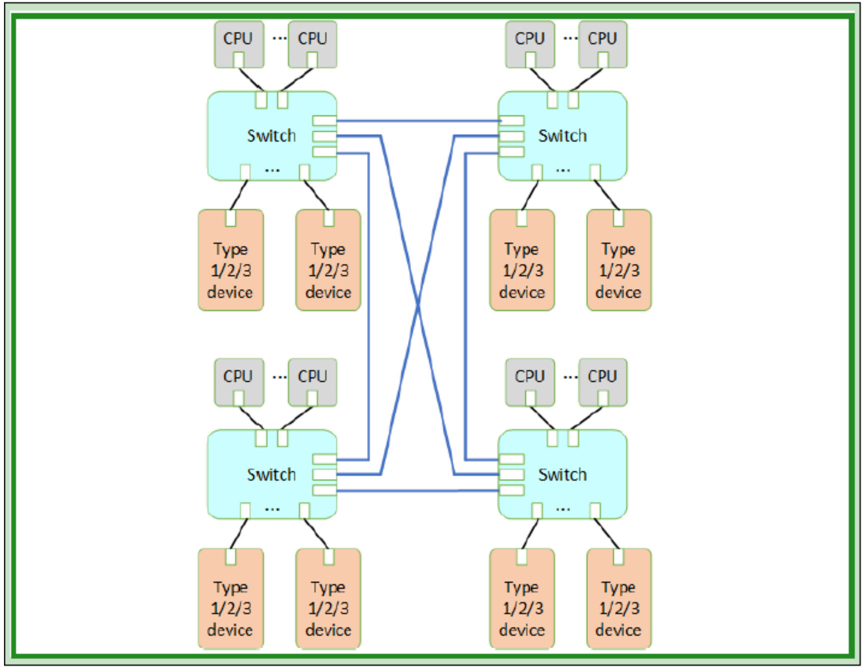
CXL 3.0: Fabric



- Not subject to traditional tree-based hierarchies
- Device communication is no longer restricted to a single coherent HPA "domain"
- Fabric switches and fabric attached memory devices are defined

Fabric Variations





Fabric Switches

- Expand from node level switch hierarchies to rack level and beyond interconnects
- Port Based Routing (PBR) switches enable cross domain access
- All traffic within the fabric is routed by PBR-ID, not address, allowing inter-switch links (ISLs) to carry traffic from multiple VHs without worrying about address collisions
- 12 bit Source Port & Destination Port IDs (SPIDs & DPIDs) enable expansion up to 4096 hosts or devices
- Global Fabric Attached Memory (GFAM) devices defined enabling enormous memory pools.
- Example: 2000+ hosts accessing a memory pool of 2000+ GFAM devices (GFDs)

CXL 3.0 Specification

Compute Express Link™

computeexpresslink.org

Thank you!

EMPOWERING OPEN.

OCP GLOBAL SUMMIT

OCTOBER 18-20, 2022
SAN JOSE, CA