



Flash Memory Summit

Memory Pooling Solution

H3 Platform
Brian Pan

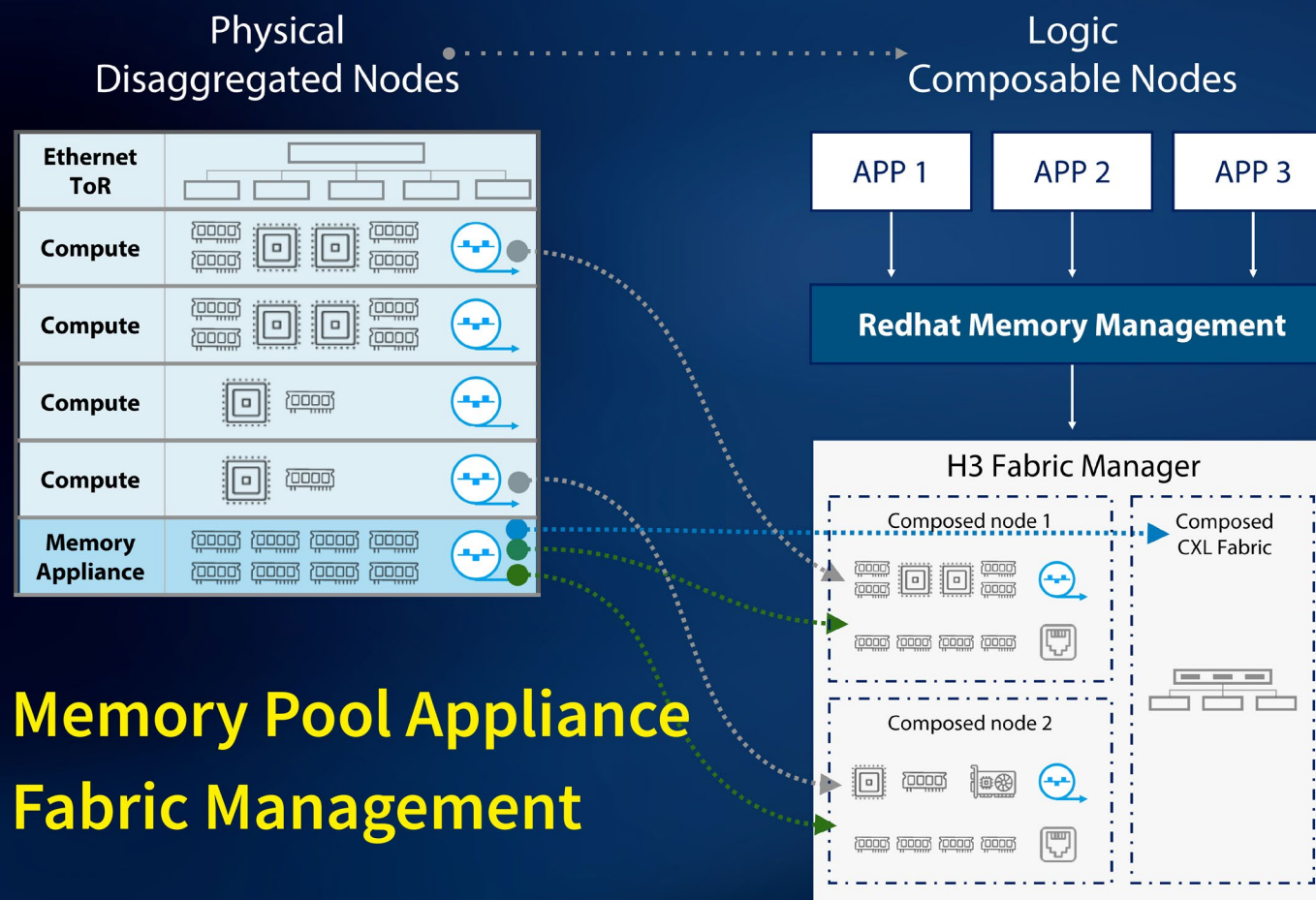


Agenda

- CXL Solution Architecture
- Key Specification and Features
- Graphic User Interface
- Fabric Management API
- Testing Results



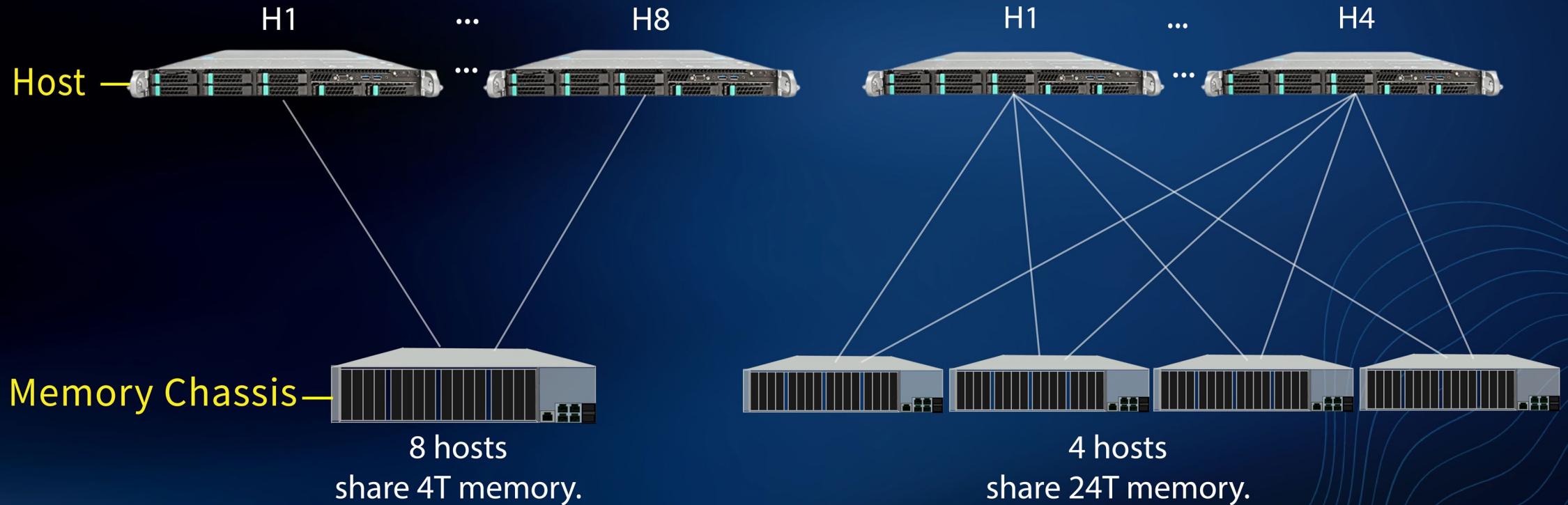
Memory Pooling Solution





CXL Memory Pooling/Sharing Usage

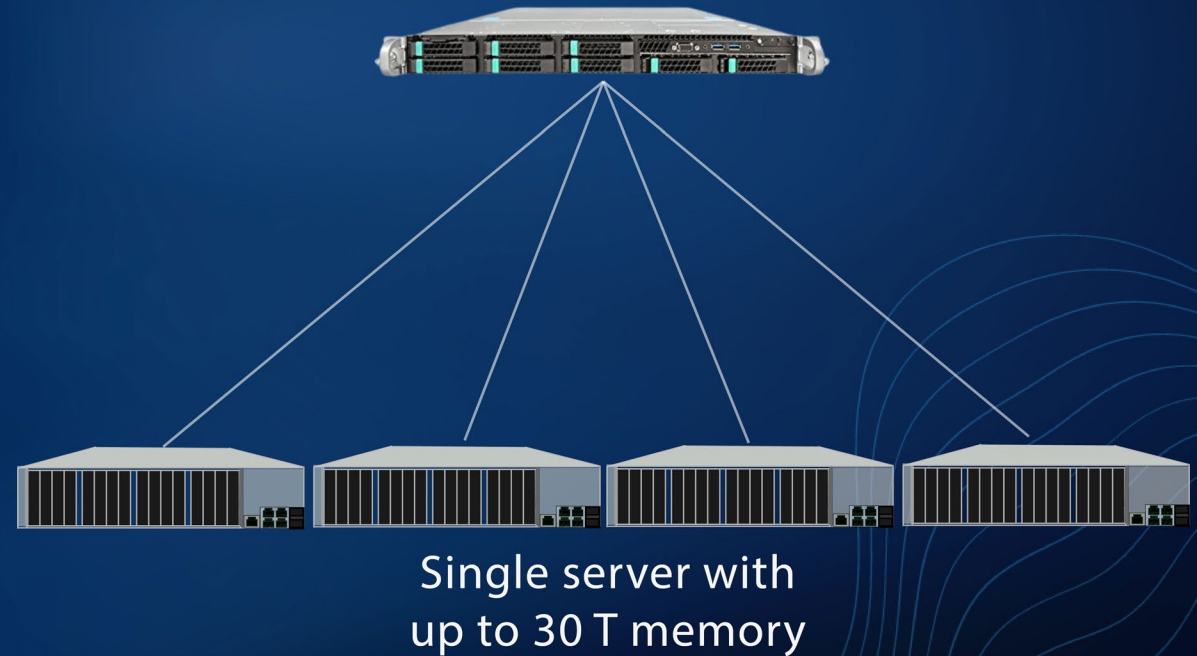
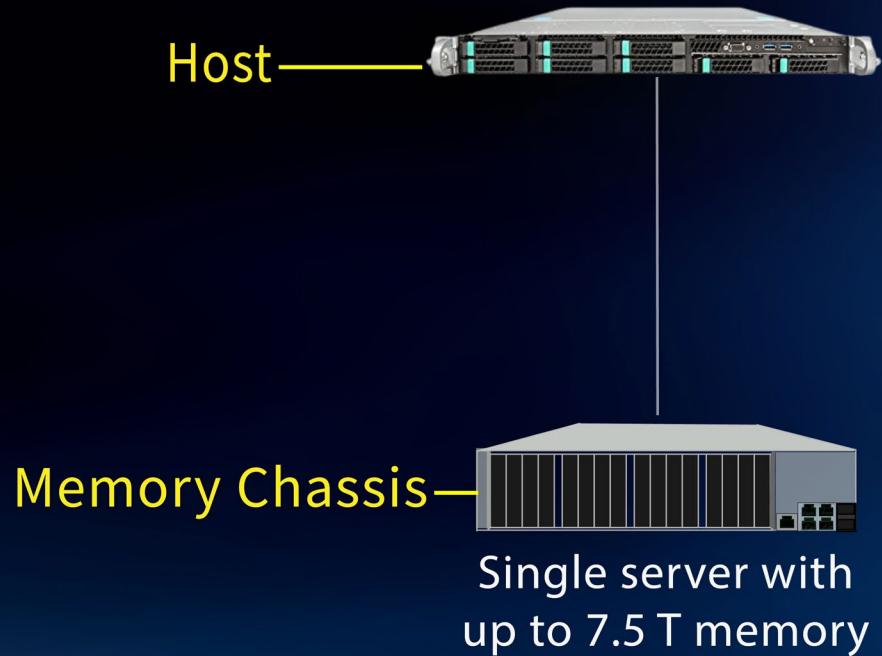
8 Hosts





CXL Memory Expansion Usage

Single Host





Software Architecture of Fabric Manager

H3 Platform GUI

User Account Management / Enclosure Information /
Healthy Status / CXL Memory Resource Management

H3 Platform BMC

Memory Device Information
Enclosure Sensors Information
XCONN Apollo Controller Information
Console Mode Only

H3 Platform RESTful API

GET/POST/PATCH/DEL/CXL/Hosts/xxx
/CXL/Devices/xxx/CXL/MDs/xxx

H3 Platform Daemon 2

Fan Speed / Power / GPIO / LED / Memory Device
Information / Enclosure Sensors

H3 Platform Daemon 1

XCONN_CMD_REGISTER / XCONN_CMD_QUERY /
XCONN_CMD_ALLOC / XCONN_CMD_ATTACH / XCONN_CMD_FREE /
XCONN_CMD_CLEAR / XCONN_CMD_EXIT /
H3_CMD_PORT_INGRESS_QUERY / H3_CMD_PORT_EGRESS_QUERY /
H3_CMD_PORT_ERROR_QUERY

Apollo Management over
UART

Apollo Management over PCIe
Mgmt. EP PCIe IDs: 1F16-C500 or 144D-A808

Memory Device

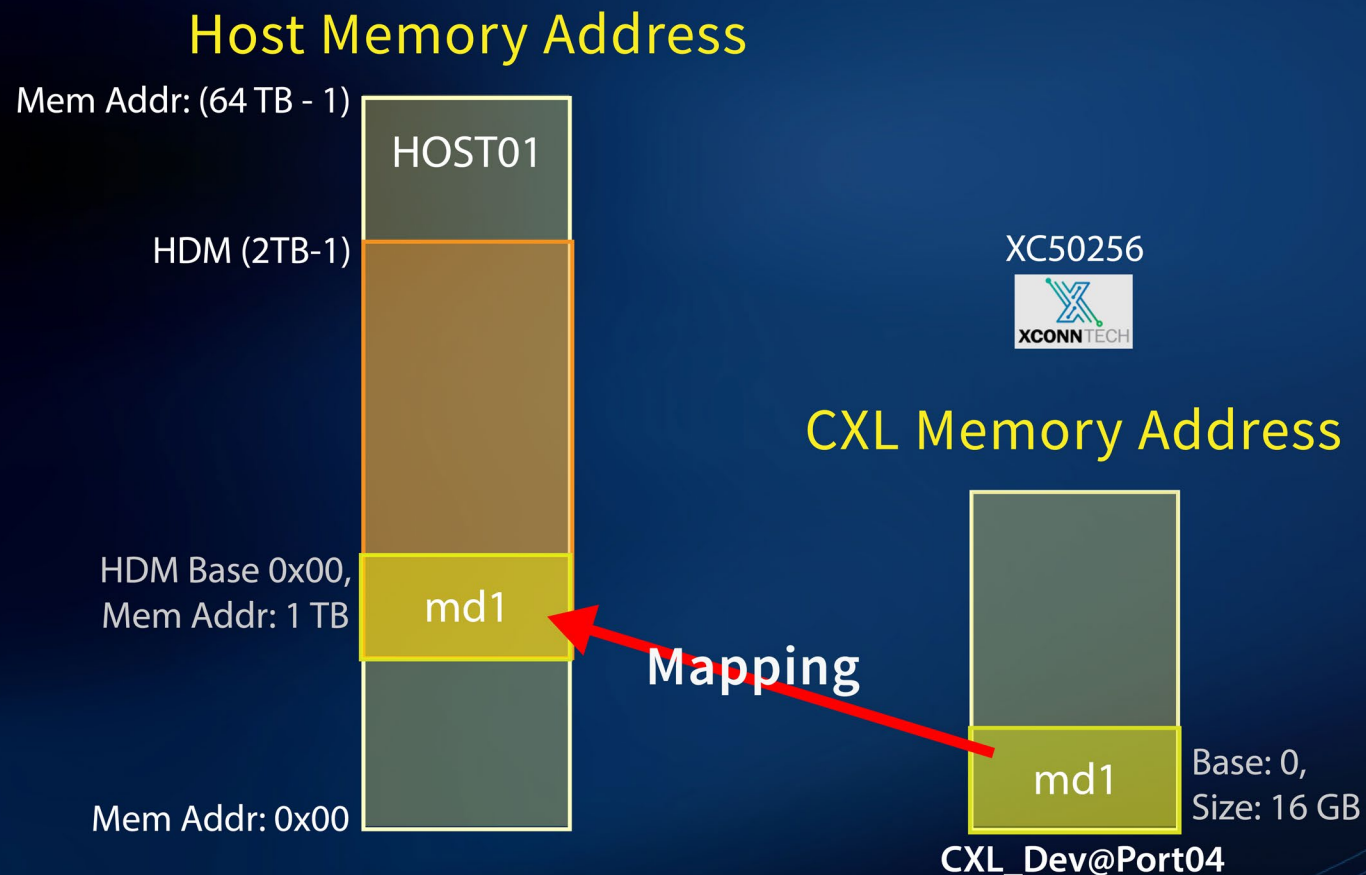
Inband/Outband IF

Apollo Switch Core



CXL Memory Pooling

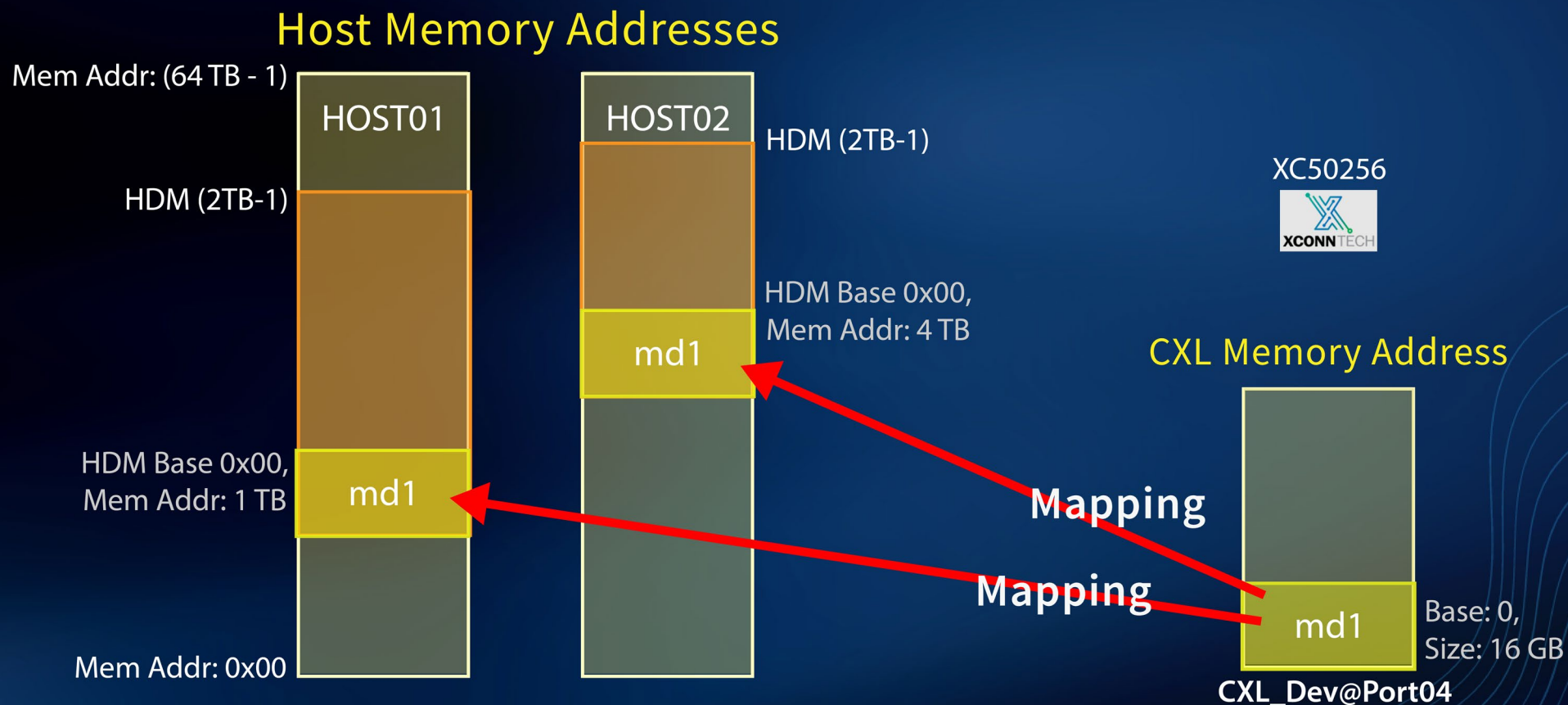
Mapping Global to Local Memory Address





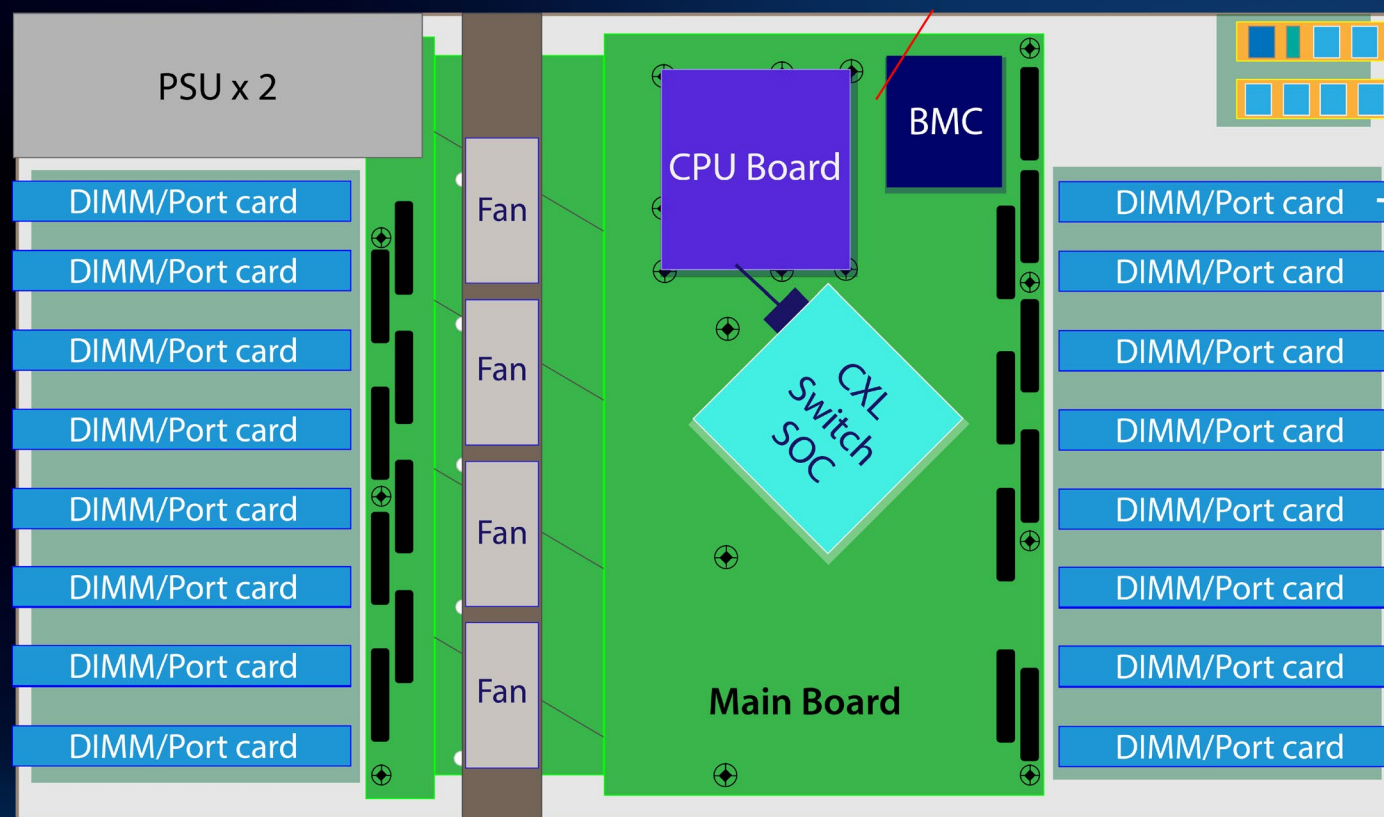
CXL Memory Sharing

Mapping Global to Multiple Local Memory Address





Flexible Port Configuration of CXL Slots





Management User Interface

Host configuration

Overview Resource Management **Host** Device

Host overview

Host List

- HOST01 3 / 16
Entry 0 MD Q'ty
Link (Curr/Max)
Gen5x16/Gen5x16
- HOST02 5 / 16
Entry 1 MD Q'ty
Link (Curr/Max)
Gen5x16/Gen5x16
- HOST03 1 / 16
Entry 2 MD Q'ty
Link (Curr/Max)
Gen5x16/Gen5x16
- HOST04 0 / 16
Entry 3 MD Q'ty
Link (Curr/Max)
Gen5x16/Gen5x16

Host managed Device Memory (CXL HDM Window)

Device memory in host

Base: 8 TB HOST01 (1024 GB) Limit: 9 TB

md09 (512GB) md05 md10 (128GB)

All MD List (12 / 64)

| ID | Memory | Share/Private | Connect Hosts | NickName |
|------|--------|---------------|---------------|-----------|
| md01 | 4GB | Private | HOST02 | DEVICE_07 |
| md02 | 256GB | Private | HOST02 | DEVICE_07 |
| md03 | 252GB | Private | HOST02 | DEVICE_07 |
| md04 | 128GB | Shared | HOST03 | DEVICE_05 |
| md05 | 48GB | Shared | HOST01 | DEVICE_05 |
| md06 | 16GB | Private | HOST02 | DEVICE_05 |
| md07 | 40GB | Private | HOST02 | DEVICE_05 |
| md08 | 128GB | Private | | DEVICE_04 |

HOST01 MD List (3 / 16)

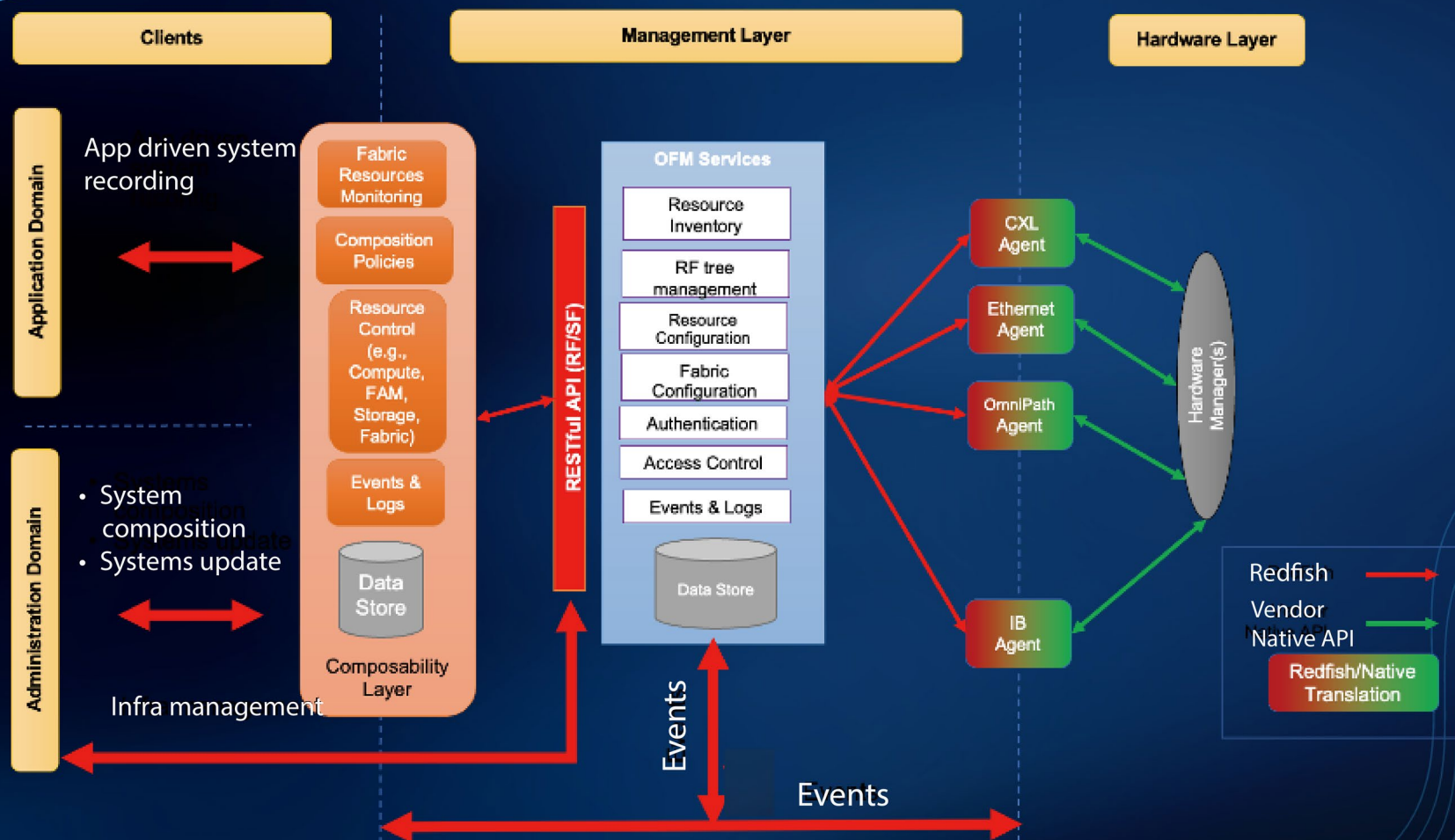
| ID | Memory | Share/Private | Connect Hosts | Detach |
|------|--------|---------------|---------------|--------|
| md05 | 48GB | Shared | HOST01 | ⊖ |
| md09 | 512GB | Private | HOST01 | ⊖ |
| md10 | 128GB | Shared | HOST01 | ⊖ |

All MD list and binded hosts

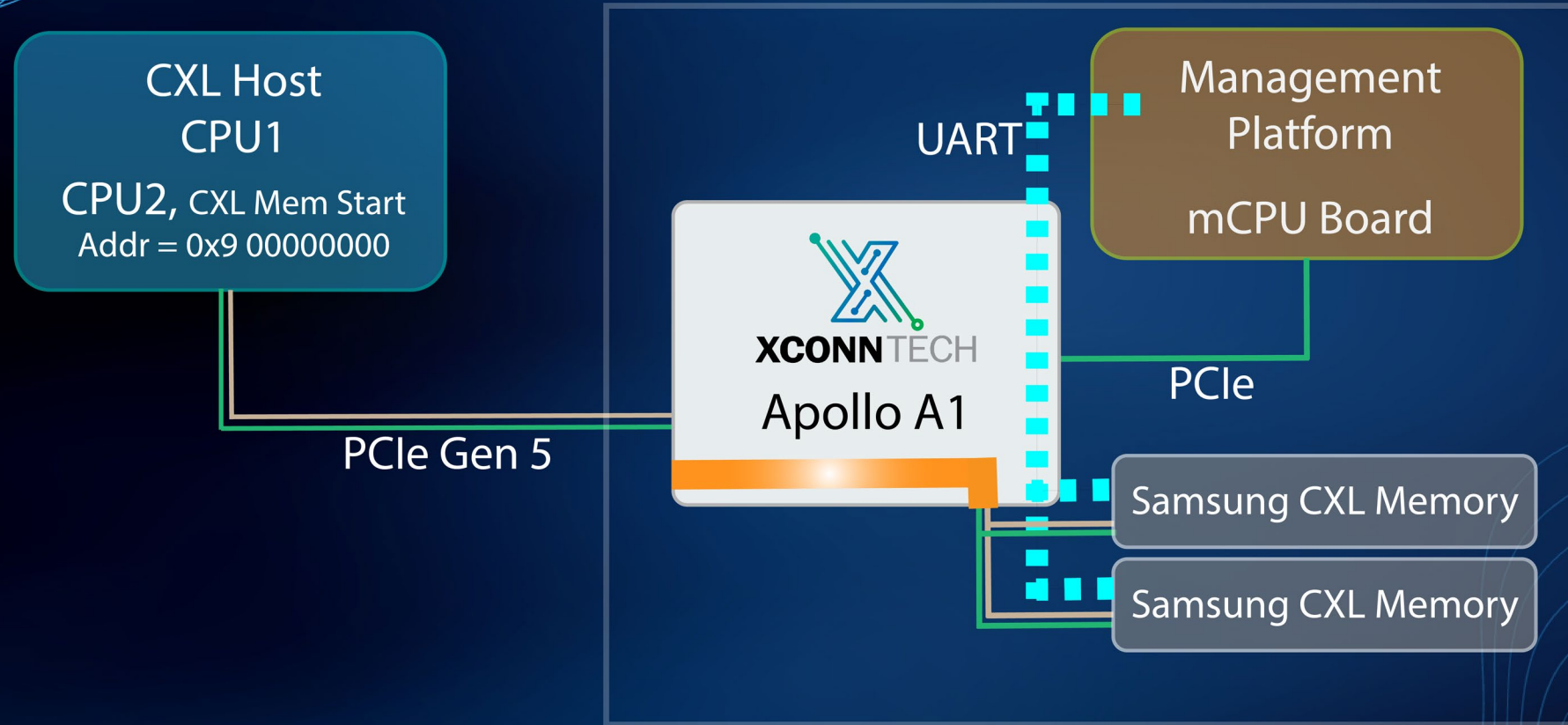
MDs belong to host



Open Fabric Management Framework



Demo System Setup



Memory Pooling Solution



Demo System Specification

Items

Server

Details

- Intel Sapphire Rapids
- Dual Xeon 8470 (2000MHz)

CXL Memory

- Samsung 128GB CXL Memory

OS

- Ubuntu 22.04.2 LTE
- Linux Kernel 6.2.0-060200-generic
- daxctl – version 72.1+
- daxio – version 1.11.1
- mlc – version 3.10



Results of Intel Memory Latency Check

| Node | 0 | 1 | 2 |
|------|----|----|----|
| 0 | 10 | 21 | 14 |
| 1 | 21 | 10 | 24 |
| 2 | 14 | 24 | 10 |

- Node 2 is the Samsung CXL memory module
- The Samsung CXL memory module is connected to the CXL port of Node 0 CPU

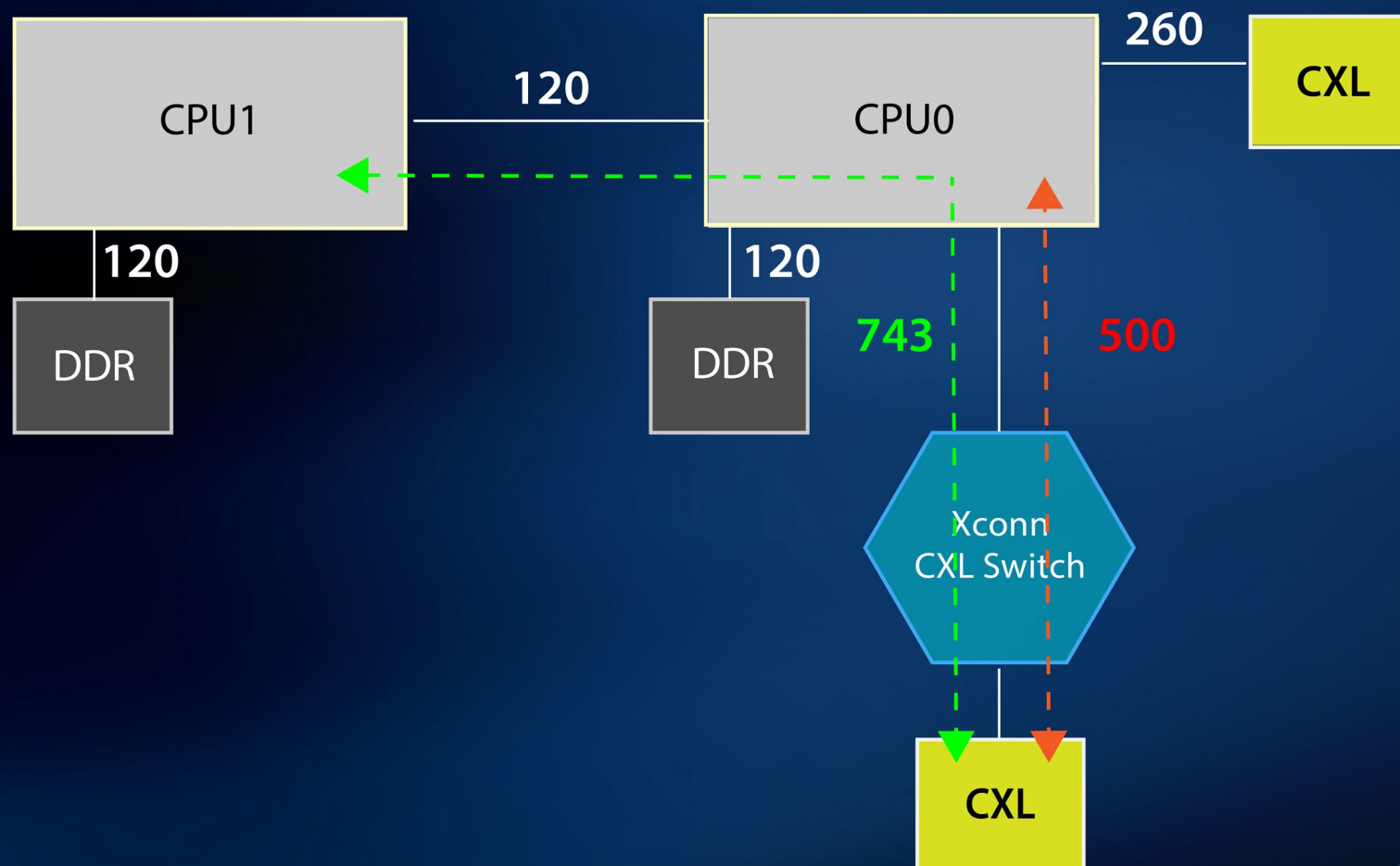


Intel MLC Results — Idle Latency (ns)

| | DDR | Direct-CXL | Switch-CXL | Note |
|--------|-----|------------|------------|------------|
| Node 1 | 120 | 260 | 500 | Close Node |
| Node 2 | 120 | 260 | 743 | Far Node |
| Diff | 0 | 140 | 240 483 | |



Intel MLC Results — Idle Latency (ns)





Result Analysis

- Latency optimization
 - CPU
 - CXL controller
 - CXL switch
 - Re-timer
- Latency breakdown
 - Will use the protocol analyzer to break down the latency in each part