FMS 2024

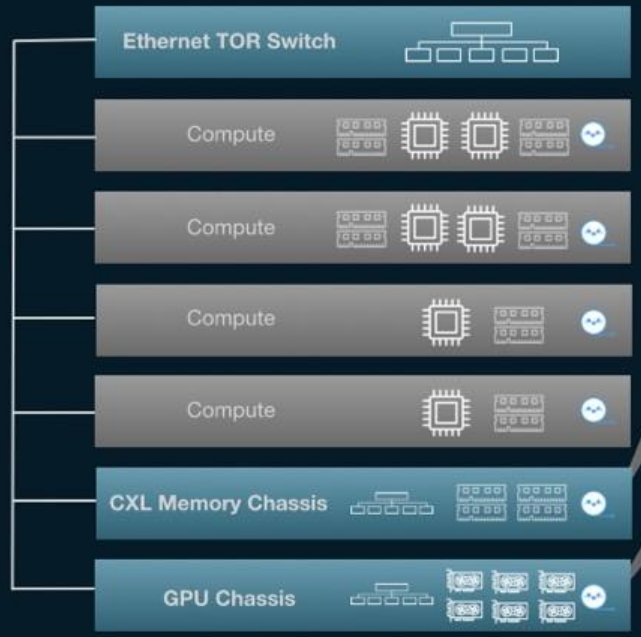
CXL Memory Sharing System Architecture and Solution Demo

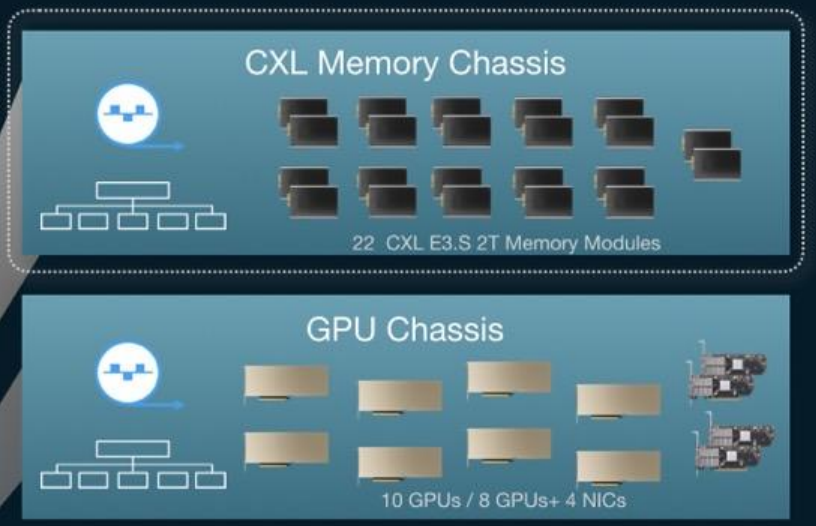
Brian Pan
H3 Platform

TABLE OF CONTENT

ARCHITECTURE COMPOSABLE PCIE SYSTEM

Physical Disaggregated Nodes





Fabric management

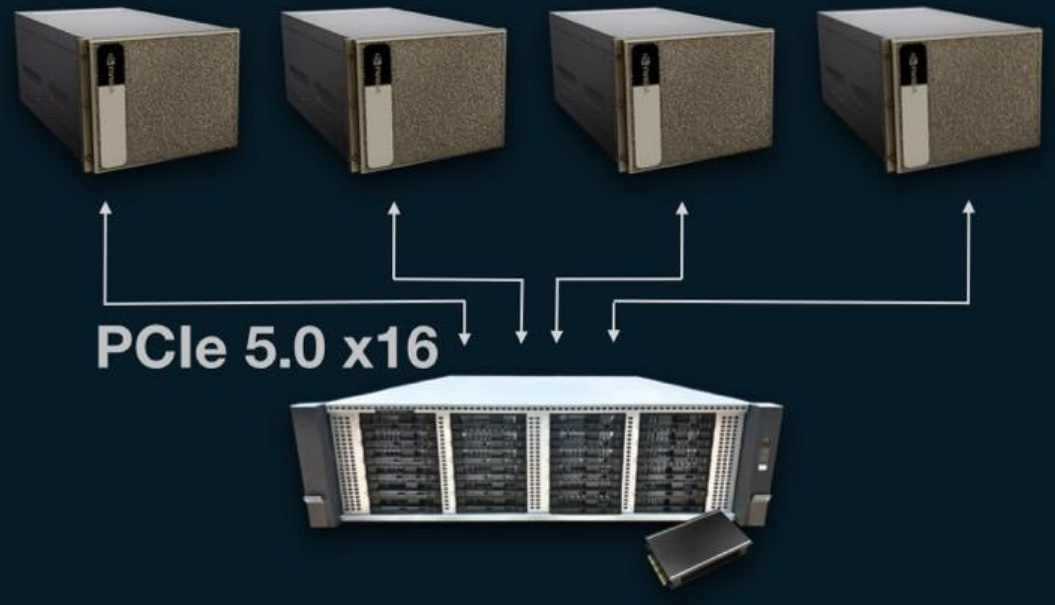
+

CXL switch

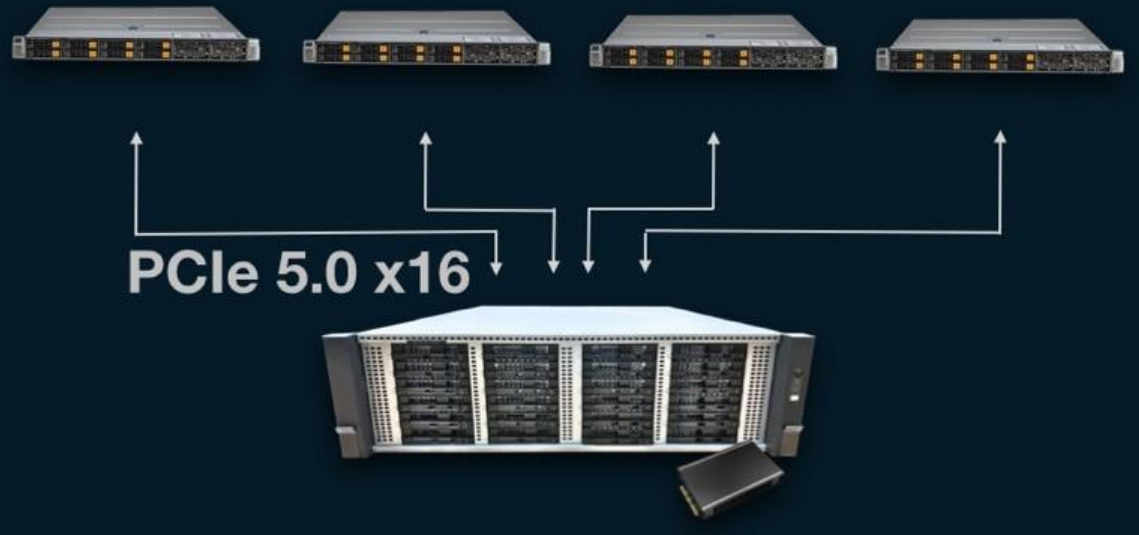
+

E3.S CXL Memory Modules

SHARING CXL MEMORY AMONG SERVERS



AI Training or Inference



In-memory Database

CHASSIS FRONT VIEW

Top Cover

Press the release button to open the top cover and to pull out the fans for replacement

Button & LEDs

Power control
Power LED
UID LED

E3.S 2T Memory Slots

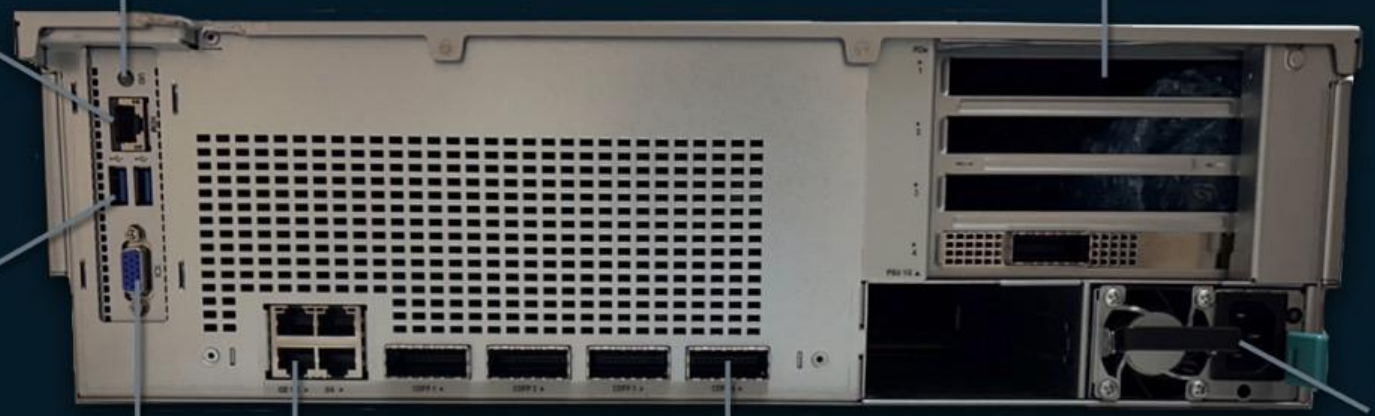
22 E3.S 2T slots
PCIe Gen5 x8

CHASSIS REAR VIEW

UID LED
Quickly locate device

Expandable PCIe Slots
Flexible configuration for double-width CXL AIC form factor cards or retainer cards

BMC Port
RJ45 Ethernet port of BMC for chassis management



USB Port
Only for H3 USB recovery and installation

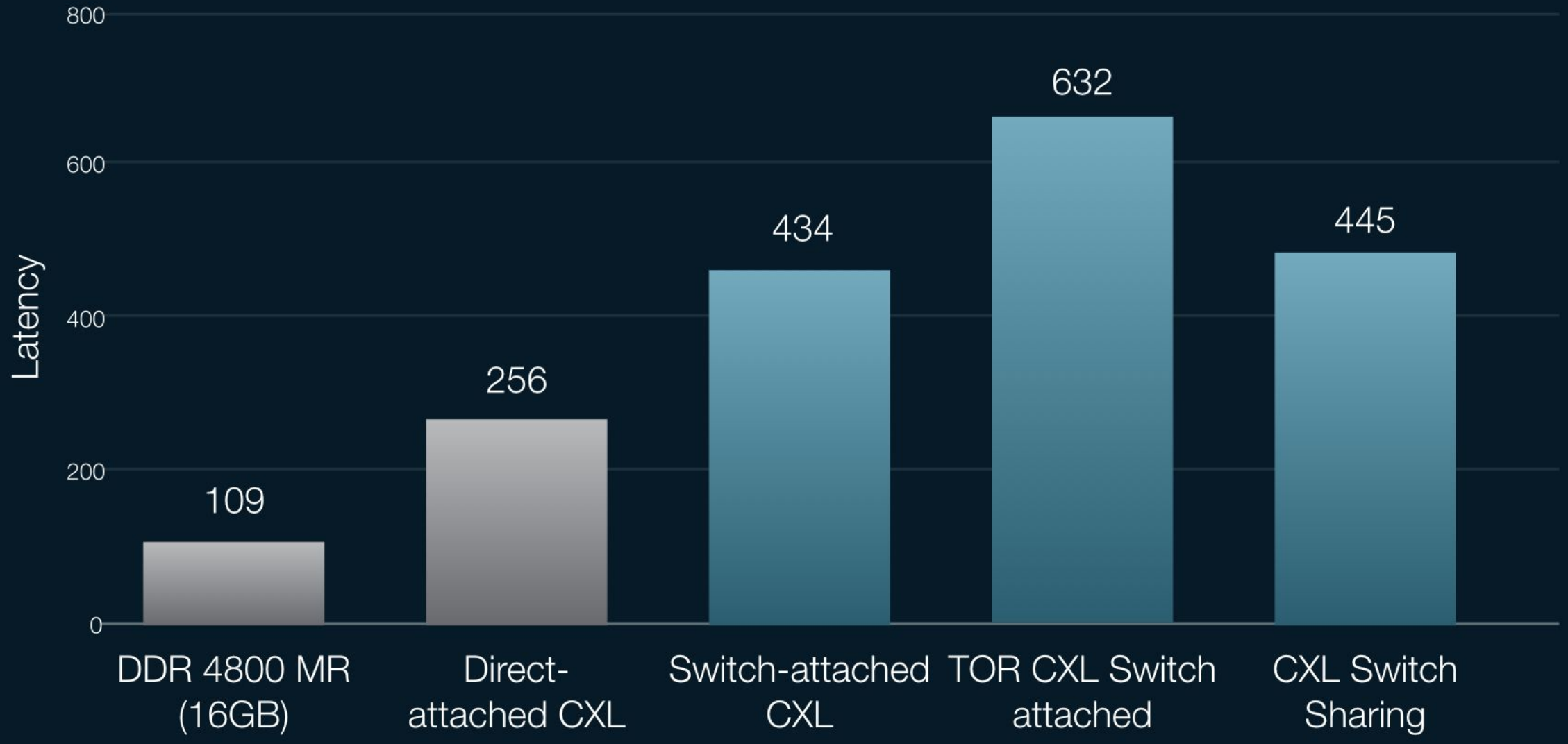
PSU
Redundant 1600W (1+1) Power Supplies

VGA Port

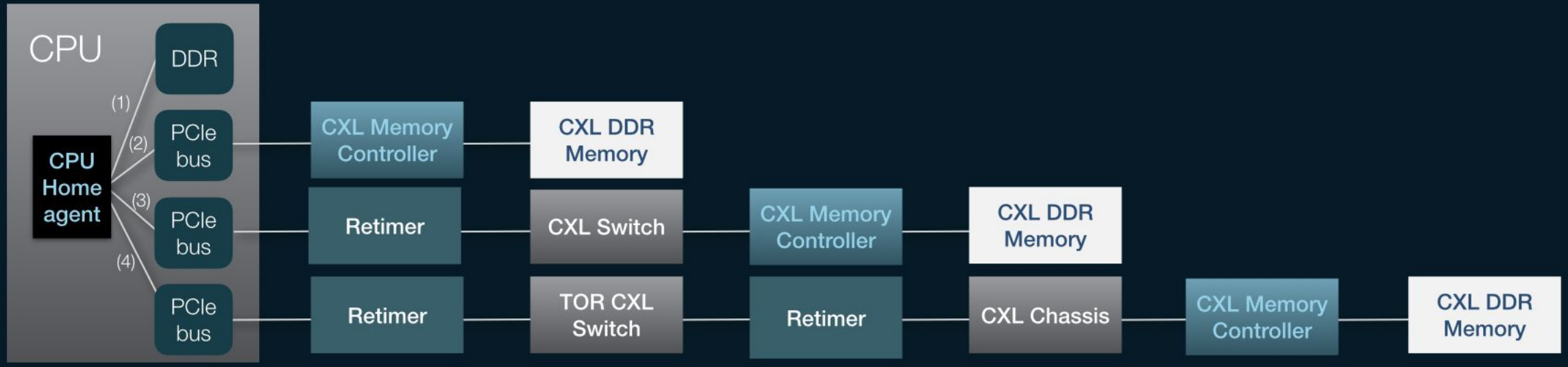
Management Port
RJ45 Ethernet port to access GUI

Host Ports
4 CDFP connectors

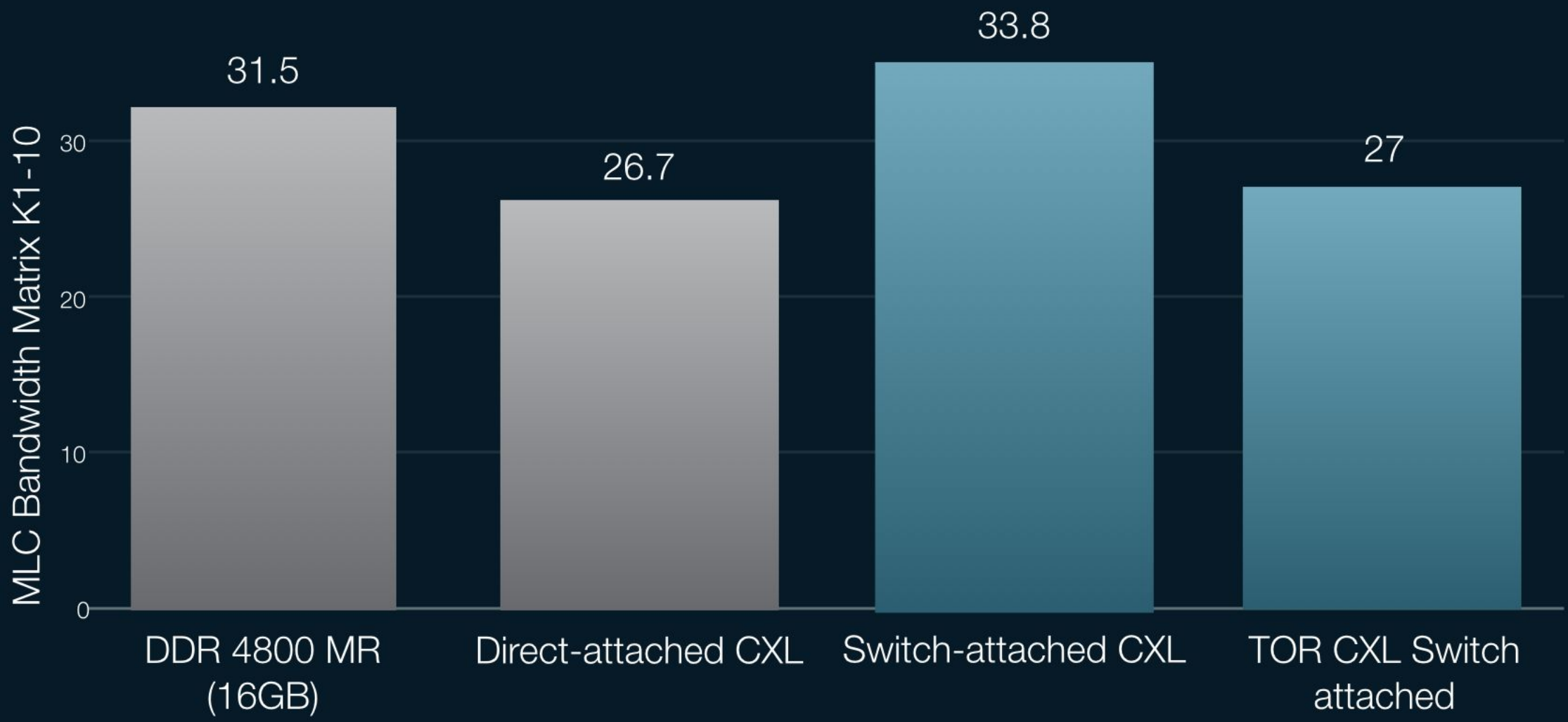
TESTING RESULT LATENCY COMPARISON OF DIFFERENT TOPOLOGY



ROUTE FROM CPU TO MEMORY



MLC BANDWIDTH OF DDR, DIRECT CXL, AND SWITCH CXL



SOFTWARE FEATURES FALCON C5022

Fabric Manager

- Composable memory sharing among hosts
- Memory surprise add and remove
- CXL port configuration to host or device ports
- Configurable host memory address
- Link capability and status

Management Interface

- Redfish®, RESTful API, GUI

TESTING SETUP

CPU

Intel EMR
Supports CXL 1.1

Linux OS

Testing 6.9.3 Now and Can Use the DAX Mode Only

DAX Driver xxx

CXL Switch

Xconn 256 Lanes CXL Switch

CXL Memory Module

Samsung

Micron

CPU PLATFORM CURRENT STATUS

CXL 1.1 Server

- CXL switch should enumerate as the the direct-attached CXL memory module for CXL 1.1 server.
- The memory address is continuous and should enter the host memory address.
- The memory is not hot-pluggable.
- The BIOS configuration is not the same in different CPU platform.

CXL 2.0 Server

- Will implement the VCS mode solution

CXL SWITCH CURRENT STATUS

CXL Switch

- Only one switch vendor now
- Still working on the latest release chipset
- Ecosystem support is critical for CXL 2.0 deployment especially CPU platform and OS

CXL MEMORY CONTROLLER AND RETIMER
CURRENT STATUS

CXL Memory Controller

- 3 major CXL memory controller released
- Different diagnostic tools for different CXL memory controller

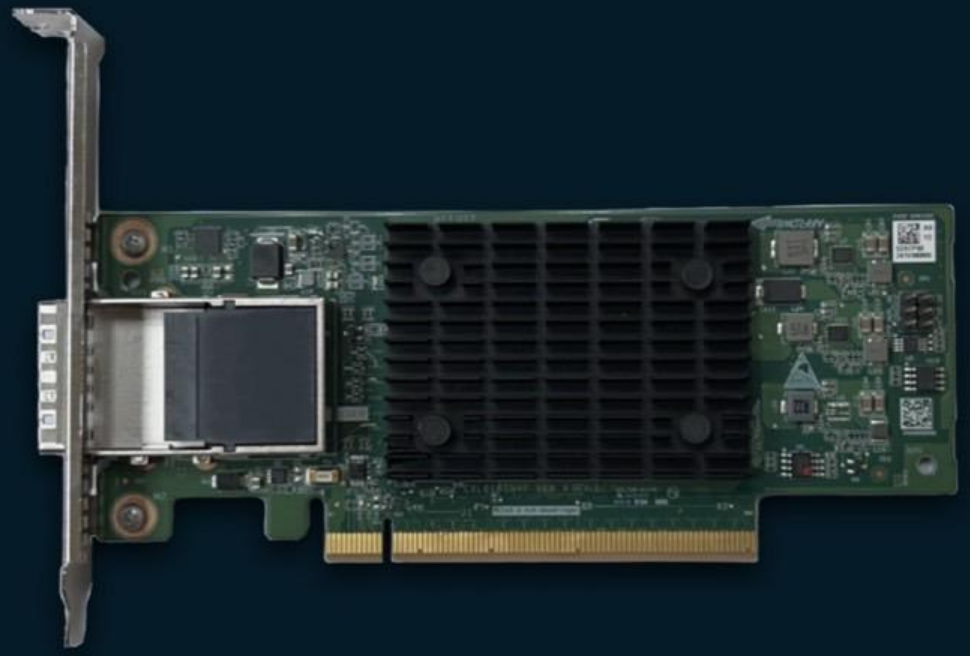
CXL Memory Modular

- E3.S is a standard form factor but the high-density E3.S is way too expensive.
- Special RAS testing tools for CXL memory controller

LESSON LEARNED RETIMER

Signal Integrity, Bifurcation, Reset, Thermal

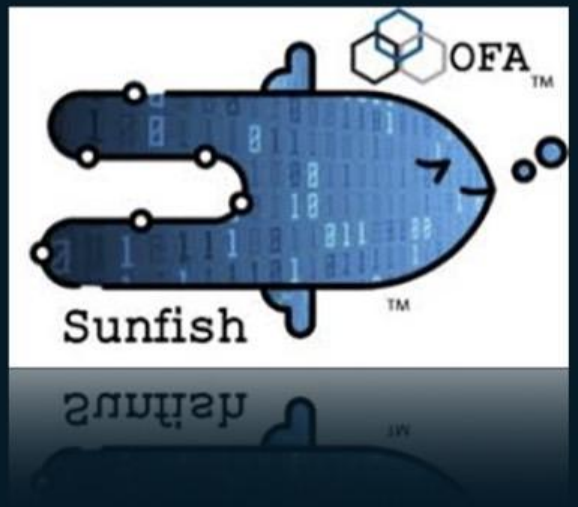
- Re-timer should test against with server slots for signal integrity
- Bifurcation, clock, and reset design
- High speed re-timer need extra cooling



LESSON LEARNED MANAGEMENT AND API

Orchestration and API

Many consortiums are working on the standard orchestration and API for the composable solutions



Management path

PCIe device management path

- Ethernet (Data or Management path)
- PCIe Inband
- I2C Out-of-band