Redfish support for Compute Express Link (CXL)

Copyright © 2023 DMTF
Agenda

• Overview of CXL
• The Chassis and System Model for CXL
  • PCIe Devices/Functions
  • CXL Logical Devices
  • Memory Domains and Memory Chunks for CXL Memory
• Modeling Local CXL Devices
  • Type 1 Devices (SmartNICs)
  • Type 2 Devices (Accelerators)
  • Type 3 Devices (Memory Buffers)
• Modeling Remote CXL Devices
  • Type 1 Devices
  • Type 2 Devices
  • Type 3 Devices
• The fabric model for all CXL Devices
  • Switches/Ports
  • Connections
Compute Express Link Overview

- Open industry standard for high bandwidth, low-latency interconnect
- Connectivity between host processor and accelerators/memory device/smart NIC
- Based on PCIe® 5.0 PHY infrastructure
  - Leverages channel, retimers, PHY, Logical, Protocols
  - Pin-to-pin compatible with PCIe connectors
  - CXL.io – I/O semantics, similar to PCIe – mandatory
  - CXL.cache – Caching semantics – optional
  - CXL.mem – Memory semantics – optional
- CXL Link supports both standard PCIe devices as well as multi-protocol CXL devices
- CXL Logical Devices allow resources partitioning and binding to different hosts

Figure above is © 2019-2022 COMPUTE EXPRESS LINK CONSORTIUM, INC.
Representative CXL Use Cases

Caching Devices / Accelerators
- **TYPE 1**
  - Processor
  - CXL
    - PROTOCOLS
      - CXL.io
      - CXL.cache
  - Accelerator
    - NIC
    - Cache
  - USAGES
    - PGAS NIC
    - NIC atomics

Accelerators with Memory
- **TYPE 2**
  - Processor
  - CXL
    - PROTOCOLS
      - CXL.io
      - CXL.cache
      - CXL.mem
  - Accelerator
    - HBM
    - Cache
  - USAGES
    - GP GPU
    - Dense computation

Memory Buffers
- **TYPE 3**
  - Processor
  - CXL
    - PROTOCOLS
      - CXL.io
      - CXL.mem
  - Memory Buffer
    - USAGES
    - Memory BW expansion
    - Memory capacity expansion
    - Storage class memory

Figure above is © 2019-2022 COMPUTE EXPRESS LINK CONSORTIUM, INC. www.dmtf.org
CXL Device Base Redfish Model

- CXL management model built on top of Redfish PCIe model
- CXL Device can provide standard PCIe Functions as well as extended functions supporting new CXL cache and memory semantics
- New CXL Logical Device allows partitioning of device resources and binding to different hosts through CXL switch
- PCIe Functions may be flexibly associated with CXL Logical Devices thus bound to different compute host (future functionality not defined in current specification)
- All PCIe Functions supporting CXL extensions associated with CXL Logical Device can use resources (e.g., memory) referenced by these devices
CHASSIS AND SYSTEM MODEL FOR CXL
Chassis Model for CXL Device

- CXL devices are not always bound to a specific host physically.
- CXL device objects are located in the Chassis Model.
  - Chassis describes resources within the chassis.
  - Remote accelerators on a CXL fabric can also reside in Chassis.
- PCIeDevice in Chassis describes CXL devices in the chassis.
  - Fabric Manager can create Memory Chunks to be available for assignment.
  - Assignment is done through the CXL Logical Device.
- Devices in Chassis describes a remote resource.
  - Shown here is a processor however other devices such as memory or I/O could also be represented here.
System Model for CXL

- System describes a host that is attached to a CXL fabric
  - Ports of processor describes how the system is connected to the fabric
  - 'Memory' can contain remote memory that describes memory bound to the host over the fabric
- MemoryDomain in system can describe either locally attached CXL memory or remote CXL memory
- MemoryChunk in system describes a section of address space backed by local or remote memory
  - Can be interleaved or a contiguous range across local or remote devices
- Memory describes a logical memory device that represents the memory that is assigned to this host over the CXL Fabric
Attached directly to Computer System

LOCAL CXL DEVICE MODELS
Local CXL Type 1 Device Model

- Systems
  - CXL-System
- Chassis
  - CXL-Chassis
  - CXL1
- PCIe Devices
  - 1
- PCIe Functions
  - 1
- CXL Logical Devices
- Processors
  - CPU
- Processors
  - Member of collection
- Processors
  - PCIe Functions
  - CXL Logical Device
- Processors
  - FPGA
- FPGA
  - PCIe Function Processor
- Service Root
  - Collection resource
  - Singleton resource
  - Subordinate relationship
  - Associate relationship

www.dmtf.org
Local CXL Type 2 Device Model

- Systems
  - CXL-System
- Chassis
  - CXL-Chassis
  - CXL2
- PCIe Devices
  - 1
- CXL Logical Devices
  - 1
- Memory Domains
  - 1
- Processors
  - CPU
- Memory Chunks
  - 1
- Memory Domains
  - CXL
  - CXL2
- Memory Chunks
  - 1
- Memory
  - HBM
- Processors
  - GPU
- Memory Media Source

Collection resource
Singleton resource
Subordinate relationship
Associate relationship
Attached to Computer System through Fabric

REMOTE CXL DEVICE MODELS
Remote CXL Type 2 Device Model
EXAMPLE RESOURCES
CXL Logical Device

Remote CXL Logical Device

Remote Logical Device in PCXL3 chassis

Semantics supported by device

Current QoS settings of device

Total Size of the Logical Device

PCle Functions associated with this CXL Logical Device

Memory components exported by this device

CXL Device endpoint on the fabric
Remote Memory at Host

Showing memory in the host from a Remote source

Memory Capacity of Accessible media source

Memory provided by a remote endpoint

Location of remote Memory source

Location of memory object in System Redfish tree
Memory Chunk

```
{
  "@odata.id": "/redfish/v1/Chassis/PCXL3/MemoryDomains/1/MemoryChunks/1",
  "Id": "1",
  "Name": "Memory Chunk 1",
  "Description": "Memory chunk accessible through CXL fabric",
  "MemoryChunkSizeMiB": 4096,
  "AddressRangeType": "PMEM",
  "AddressRangeOffsetMiB": 1024,
  "RequestedOperationalState": "Online",
  "MediaLocation": "Local",
  "Links": {
    "CXLLLogicalDevices": [
      { "@odata.id": "/redfish/v1/Chassis/PCXL3/PCIeDevices/1/CXLLLogicalDevices/1" }]
  },
  "Endpoints": [
    { "@odata.id": "/redfish/v1/Fabrics/CXL/Endpoints/T3" }
  ]
}
```

- Memory range accessible by Host
- Memory chunk is in remote chassis PCXL3
- Total Size of this Memory Chunk
- Memory Type and offset within this CXL Device
- Logical Device that is exporting this Memory Chunk
- Endpoint of this device on the CXL Fabric
Additional Resources

- More details of the CXL model can be found at the CXL Public Mockup on the DMTF Website [https://redfish.dmtf.org](https://redfish.dmtf.org)
- The Fabrics Whitepaper provides details of the redfish fabrics model and example fabric types
  - [https://www.dmtf.org/sites/default/files/standards/documents/DSP2066_1.0.0.pdf](https://www.dmtf.org/sites/default/files/standards/documents/DSP2066_1.0.0.pdf)
Thank you!

For more information, visit us online at dmtf.org

Visit the Redfish Developer Hub at redfish.dmtf.org