Update on Redfish Support for Compute Express Link (CXL)

Redfish

www.dmtf.org

INFORMATION & DISCLAIMER

- The information in this presentation represents a snapshot of work in progress within the DMTF.
- This information is subject to change. The Standard Specifications remain the normative reference for all information.
- DMTF Specifications in progress are considered DMTF confidential and are being shown under existing work register agreements.
- Documents cannot be shared outside of the DMTF work group. The documents may be posted on the WG site, but not otherwise distributed without DMTF Board approval.
- Non-member comments, feedback and submissions are subject to the DMTF Patent Policy statement

Agenda

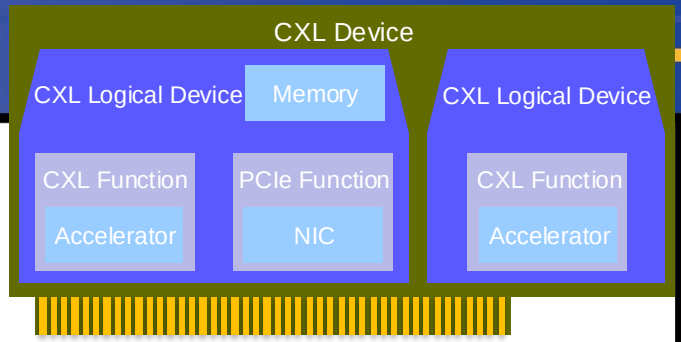
- Status of CXL in Redfish
- The Chassis and System Model for CXL
  - PCIe Devices/Functions
  - CXL Logical Devices
  - Memory Domains and Memory Chunks for CXL Memory
- Modeling Local CXL Devices
  - Type 1 Devices (SmartNICs)
  - Type 2 Devices (Accelerators)
  - Type 3 Devices (Memory Buffers)
- Modeling Remote CXL Devices
  - Type 1 Devices
  - Type 2 Devices
  - Type 3 Devices
- DCD Models
- Example Redfish Resources

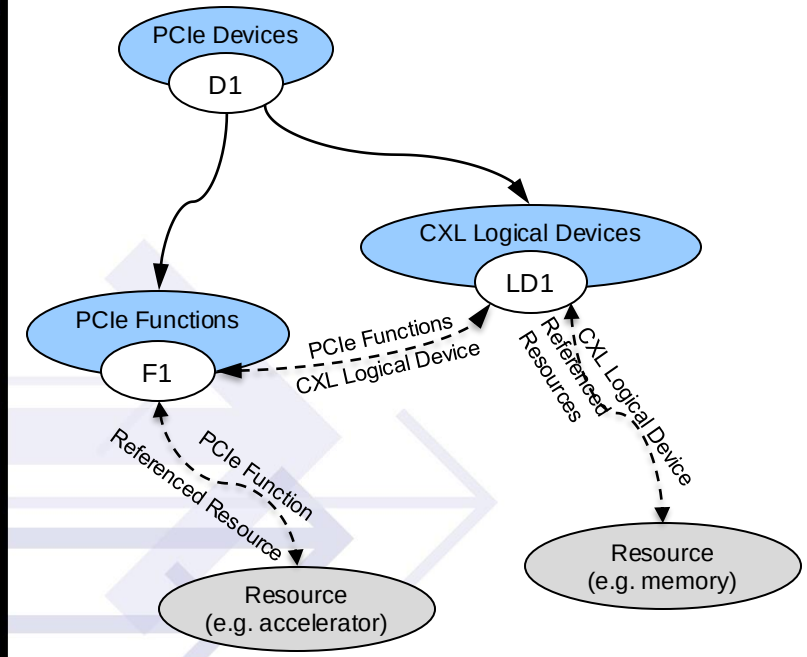
Status of CXL in Redfish

- Current State of Fabrics Model
  - Fabric model supporting CXL 2.0 are complete
  - Some CXL 3.0 models have been added
    - Basic DCD
    - CXL Switch Basic Configuration
    - Recommend start with Redfish version 2025.1
  - DSP_0288 Redfish Mapping Spec v1.2 has been released
    - Needs some review from implementers
    - Currently being updated with recent changes

- What we are working on (Fabric Task Force Work Items)
  - Dynamic Capacity Devices
  - Rosetta Stone on new capabilities in v3.1 of CXL Specification
  - Current redfish model requires a high-level orchestrator
    - Currently, no redfish construct that takes CXL CCI information for global memory and have it persist if you get rid of memory region. (No Logical Memory Region).
    - We have a WIP published for Logical Memory Regions and would like your feedback
    - DSP-IS0020_1.0.0WIP50.zip

CXL Device Base Redfish Model



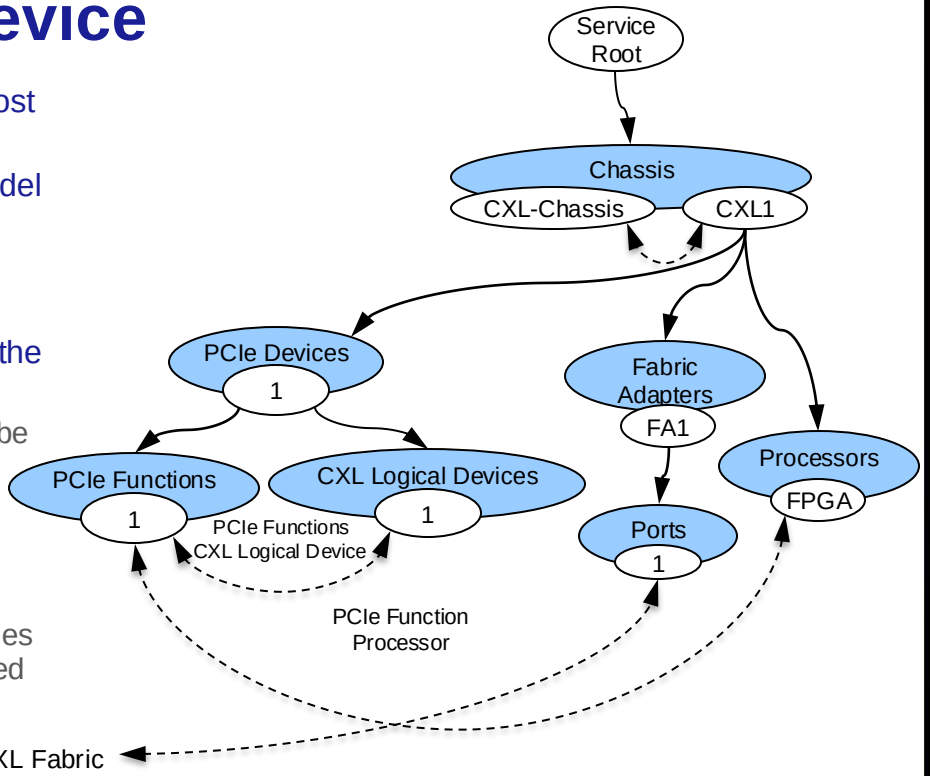


- CXL management model built on top of Redfish PCIe model
- CXL Device can provide standard PCIe Functions as well as extended functions supporting new CXL cache and memory semantics
- New CXL Logical Device allows partitioning of device resources and binding to different hosts through CXL switch
- PCIe 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

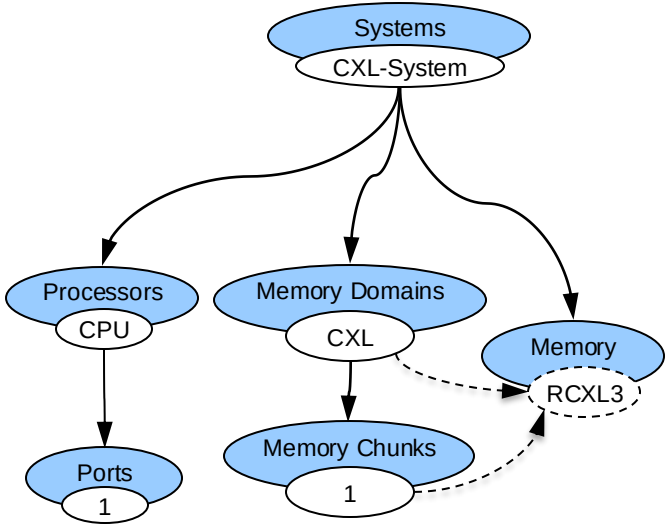
Chassis Model for CXL Device

- 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

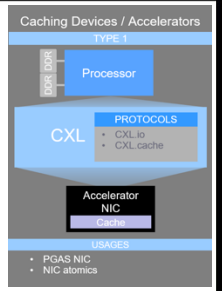
- System describes a host that is attached to a CXL fabric
  - Ports of processor describes how the system is connected to the fabric
  - ‘Memory’ can contain remote memory that describes memory bound to the host over the fabric
- MemoryDomain in system can describe either locally attached CXL memory 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

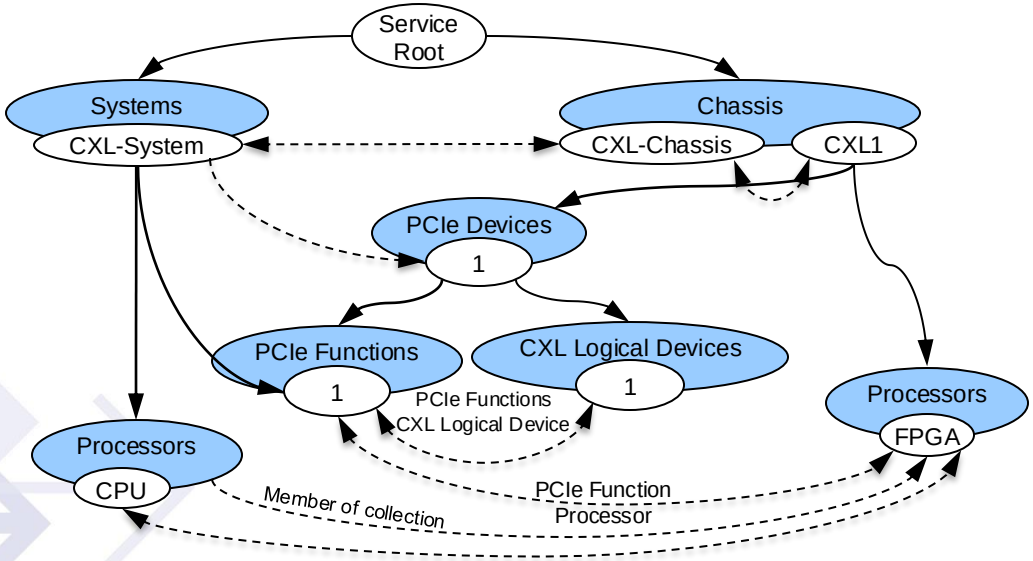


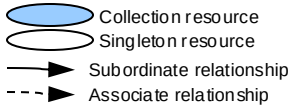
Attached directly to Computer System

LOCAL CXL DEVICE MODELS

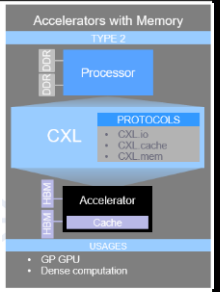
Local CXL Type 1 Device Model



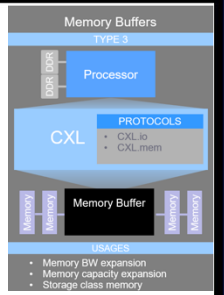




Local CXL Type 2 Device Model



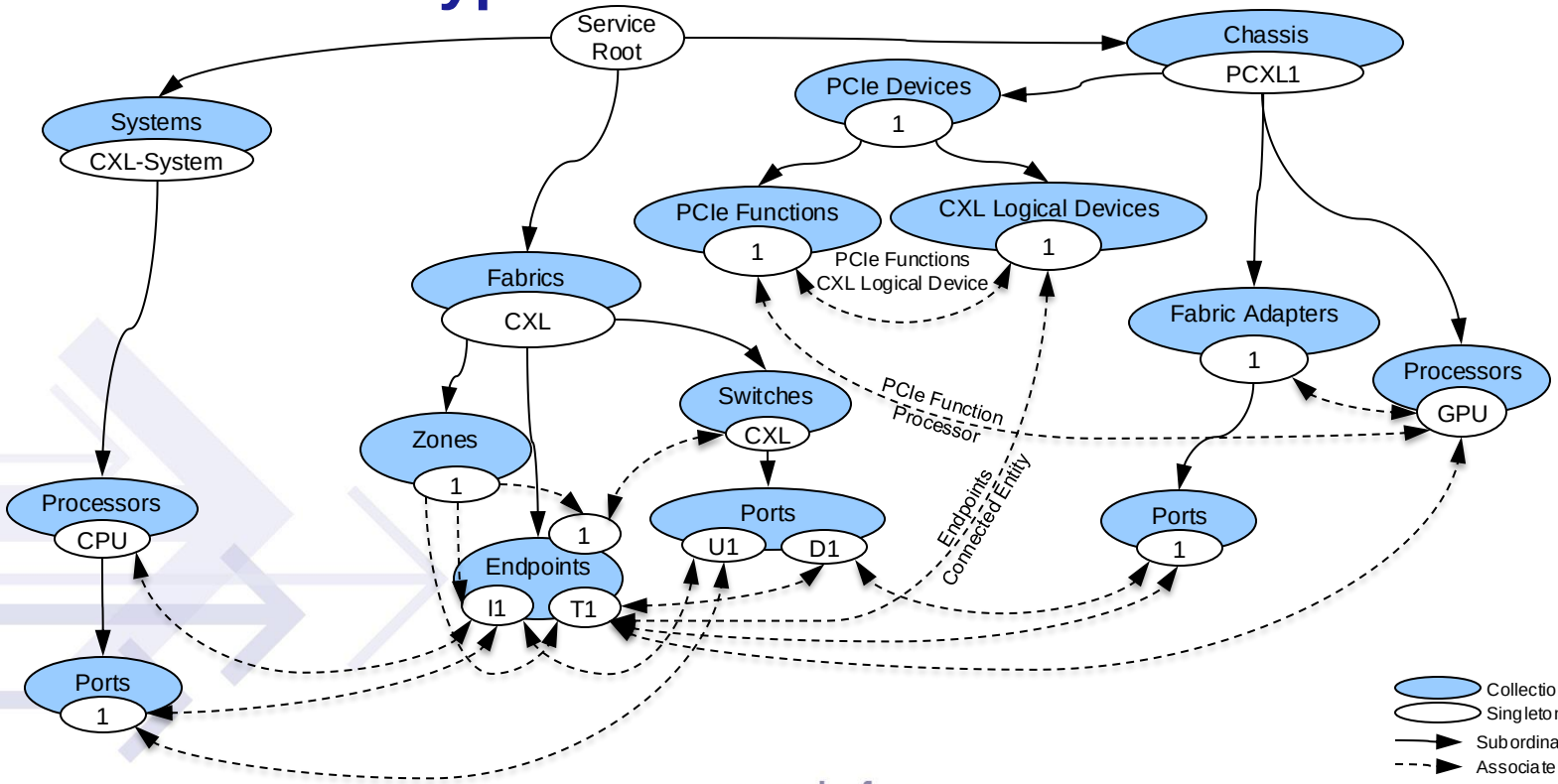
Local CXL Type 3 Device Model



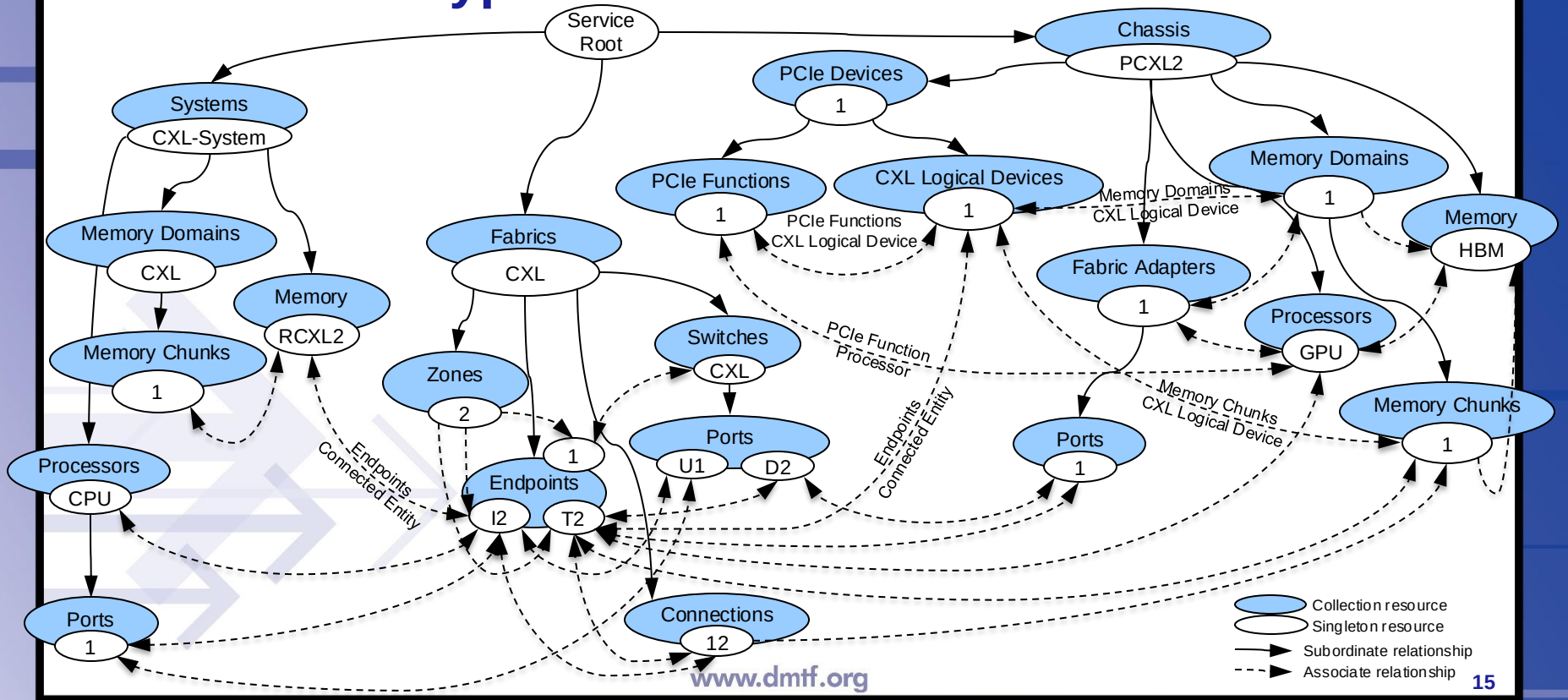
Attached to Computer System through Fabric

REMOTE CXL DEVICE MODELS

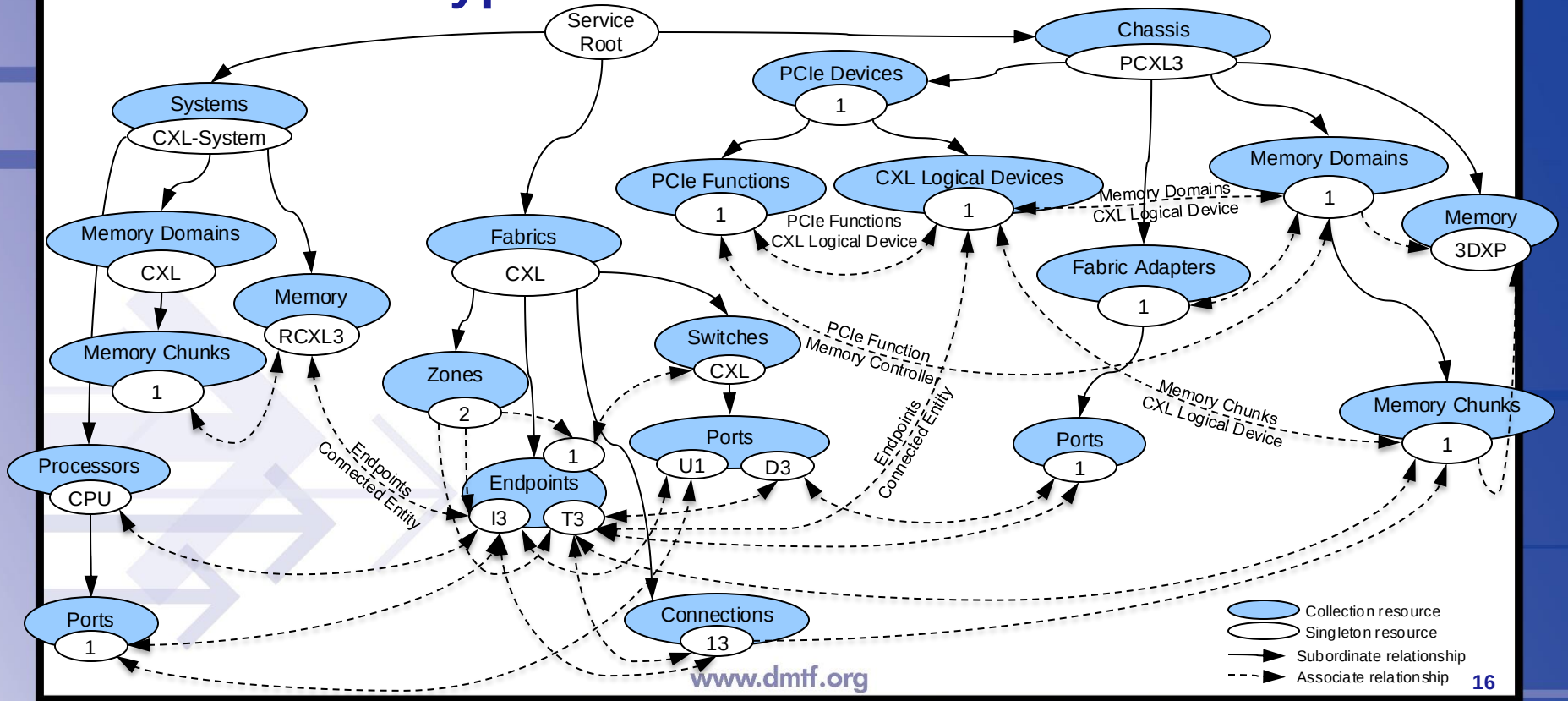
Remote CXL Type 1 Device Model



Remote CXL Type 2 Device Model



Remote CXL Type 3 Device Model

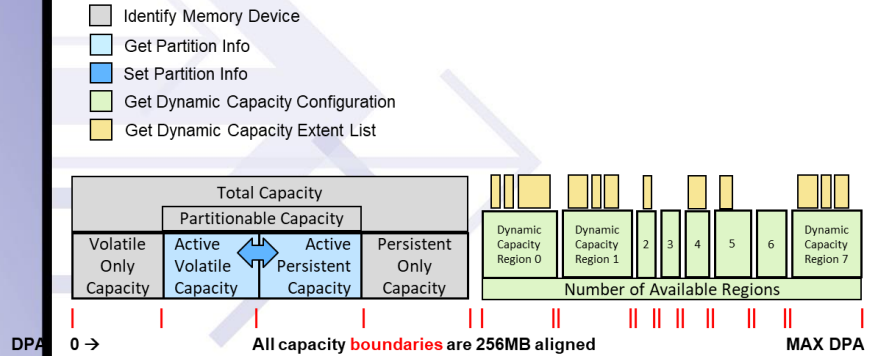


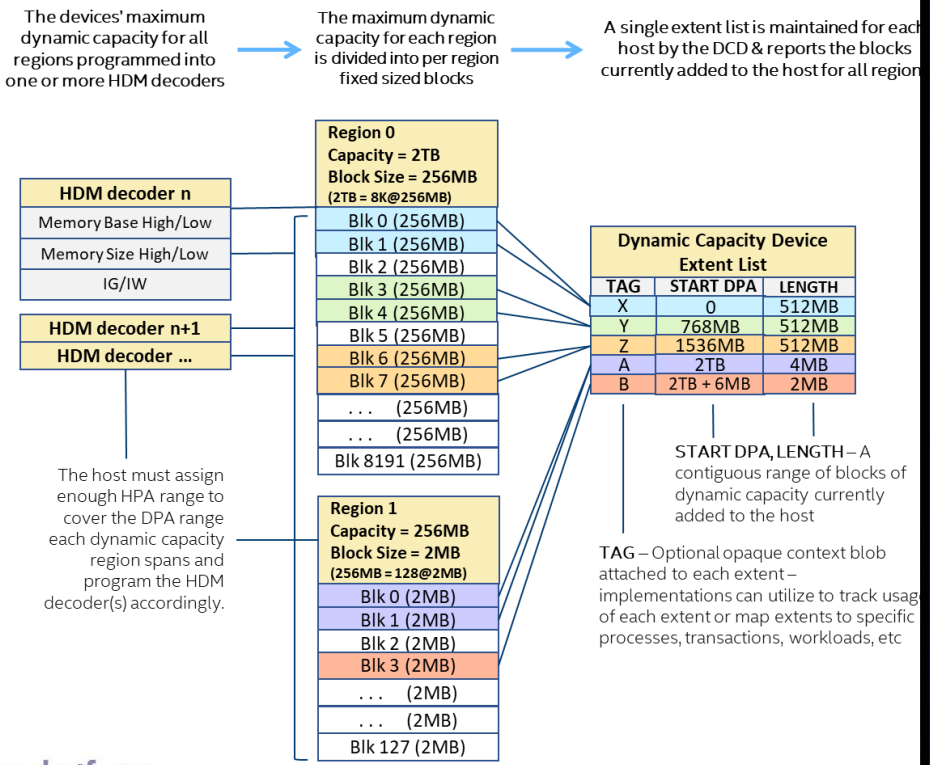
Switch attached

DYNAMIC CAPACITY DEVICES

CXL Dynamic Capacity Device

- Dynamic Capacity is a feature of a CXL memory device that allows memory capacity to change dynamically without the need for resetting the device
- DCD controls the allocation of these DC blocks to the host which requires appropriated management APIs to be defined
- CXL 3.0 Specification defines set of management commands for DCD management which needs to have counterparts in the Redfish CXL management model

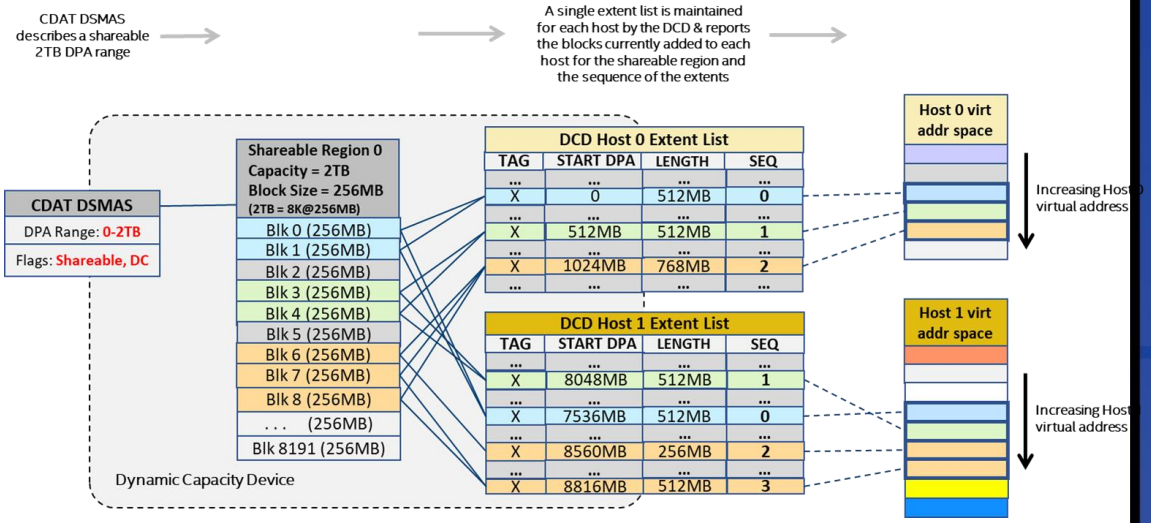




Source: CXL 3.0 Specification

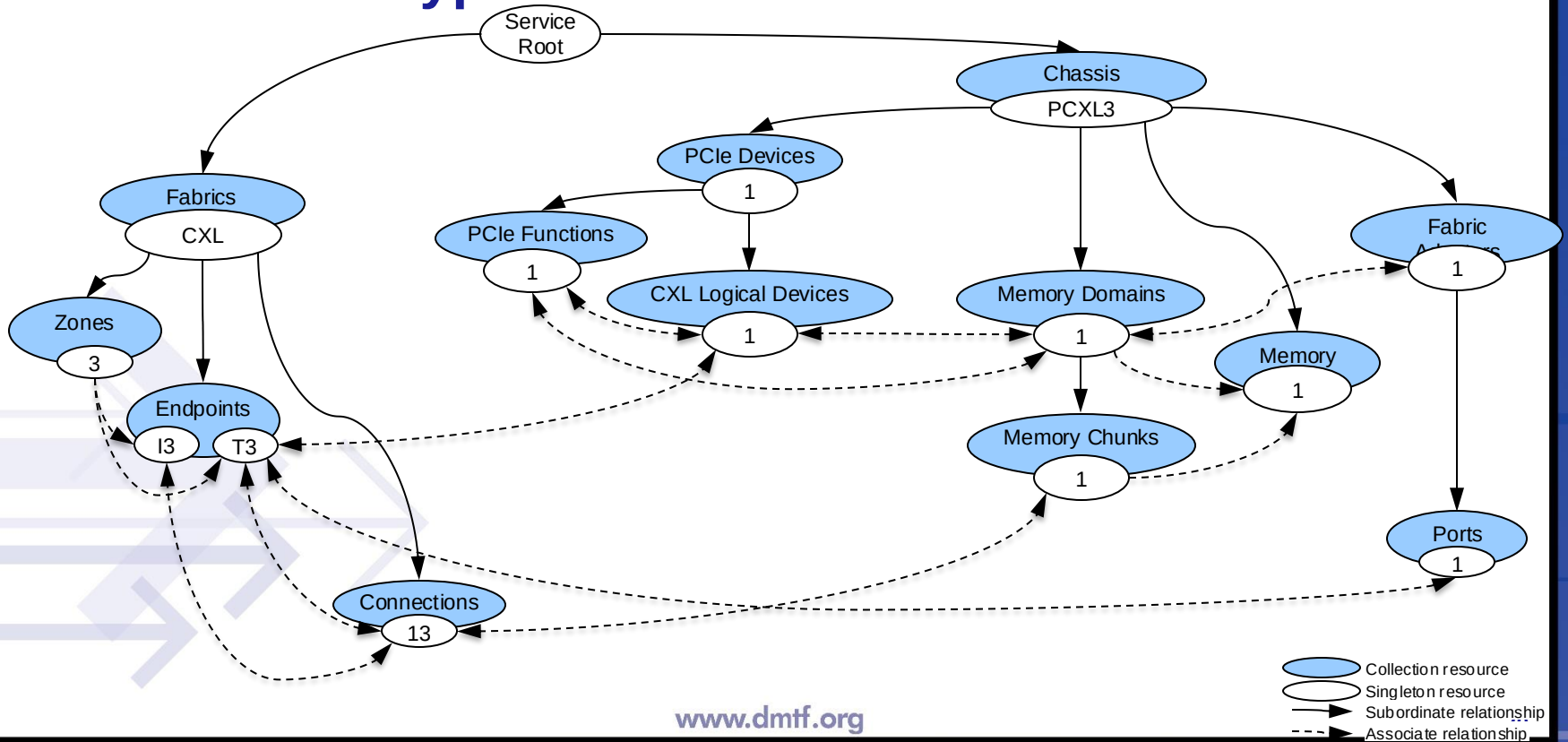
Redfish CXL Model Changes

- New resource to model memory region
  - Memory regions are defined per host so they should be subordinates of the CXL Logical Device
  - Physical memory is provided to memory region by multiple memory chunks
  - Memory region capacity is exposed to the host using extent list
- Resources updated
  - PCIe Device
  - CXL Logical Device
  - Memory
  - Memory Chunks
  - Connection

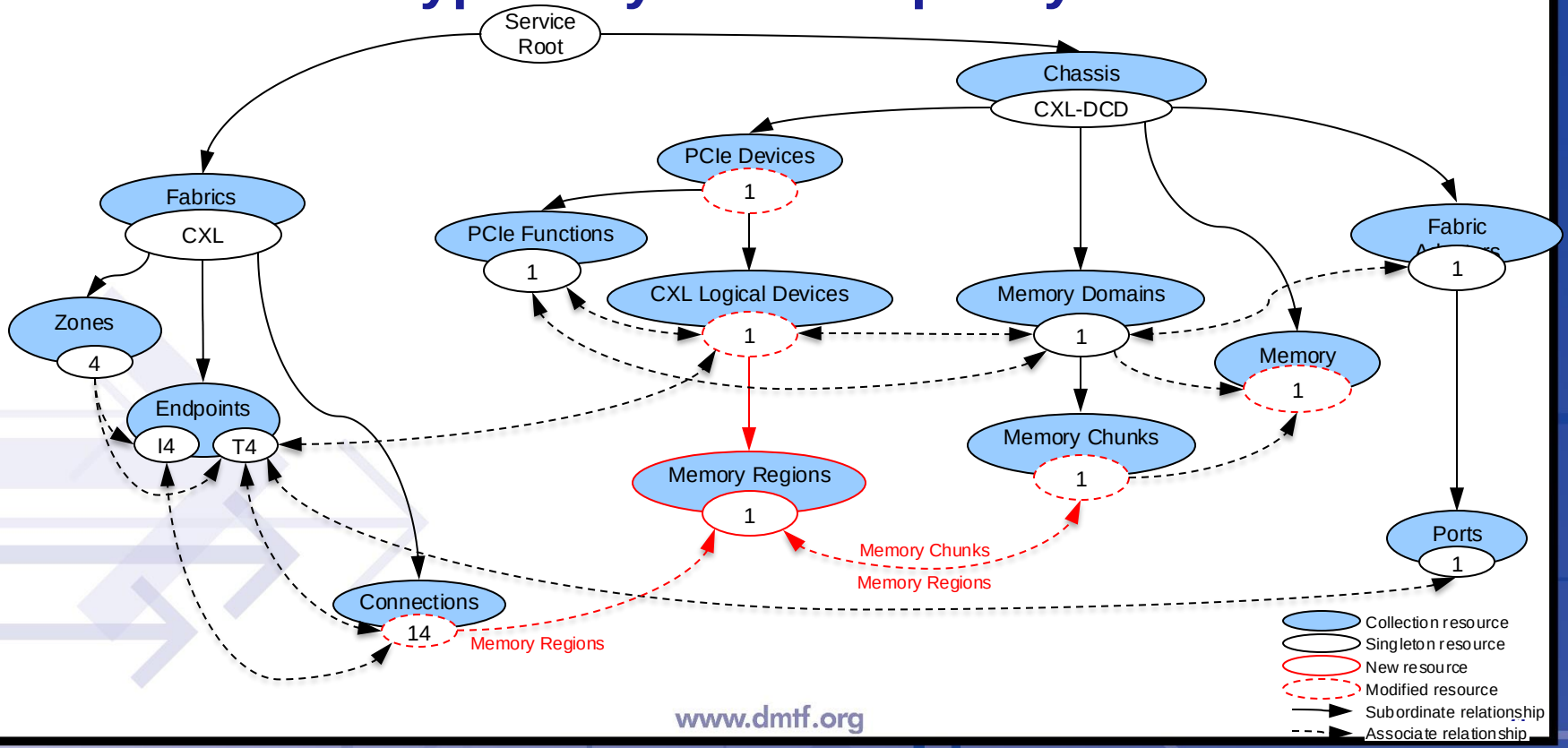


Source: CXL 3.0 Specification

Remote CXL Type 3 Current Model



Remote CXL Type 3 Dynamic Capacity Device Model



EXAMPLE RESOURCES

CXL Logical Device

Remote CXL Logical Device

```
{
    "@odata.id": "/redfish/v1/Chassis/PCXL3/PCIeDevices/1/CXLLogicalDevices/1",
    "Name": "CXL Logical Device Type 3",
    "Description": "CXL Logical Device Type 3 accessible through CXL fabric",
    "SemanticsSupported": [
        "CXLio",
        "CXLmem"
    ],
    "QoS": {
        "AllocatedBandwidth": 64,
        "LimitPercent": 85
    },
    "MemorySizeMiB": 16384,
    "Links": {
        "PCIeFunctions": [
            { "@odata.id": "/redfish/v1/Chassis/PCXL3/PCIeDevices/1/PCIeFunctions/1" }
        ],
        "MemoryChunks": [
            { "@odata.id": "/redfish/v1/Chassis/PCXL3/MemoryDomains/1/MemoryChunks/1" }
     ],
        "MemoryDomains": [
            { "@odata.id": "/redfish/v1/Chassis/PCXL3/MemoryDomains/1" }
        ],
        "Endpoints": [
            { "@odata.id": "/redfish/v1/Fabrics/CXL/Endpoints/T3" }
        ]
```

Remote Logical Device in PCXL3 chassis

Semantics supported by device

Current QoS settings of device

Total Size of the Logical Device

PCIe 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

```
{   "Id": "RCXL3",
    "Name": "CXL Device memory",
    "Description": "Remote CXL device memory",
    "CapacityMiB": 4096,
    "MemoryDeviceType": "Logical",
    "Links": {
        "Endpoints": [
            {  "@odata.id": "/redfish/v1/Fabrics/CXL/Endpoints/I3" }
        ],
        "MemoryMediaSources": [
            { "@odata.id": "/redfish/v1/Chassis/PCXL3/MemoryDomains/1/MemoryChunks/1" }
        ]
    },
    "@odata.id": "/redfish/v1/Systems/CXL-System/Memory/RCXL3",
    …
```

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

Memory range accessible by Host

```
{
    "@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,
    "MediaLocation": "Local",
    "RequestedOperationalState": "Online",
    "Links": {
        "CXLLogicalDevices": [
            { "@odata.id": "/redfish/v1/Chassis/PCXL3/PCIeDevices/1/CXLLogicalDevices/1" }
        ],
        "Endpoints": [
            { "@odata.id": "/redfish/v1/Fabrics/CXL/Endpoints/T3" }
        ]
    },
    …
```

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

DCD EXAMPLES

Memory Region Schema

- New schema defining Memory Region with following properties
  - Region Number
  - Region Type
    - **Static** – static memory region
    - **Dynamic** – dynamic memory region supporting dynamic capacity
  - Region Base offset (DPA addressable)
  - Region Size
  - Flag indicating shareable region
  - Flag indicating memory needs to be sanitized during deallocation
  - Flag indicating Non-Volatile memory region
  - Flag indicating support for hardware managed coherency
  - Region memory block size
  - Extents Count (Dynamic Memory Region only)
  - Array of Memory Extents (Dynamic Memory Region only)
    - Memory Extent Offset (DPA addressable)
    - Memory Extent Size
    - Extent user-defined Tag
    - Host Extent Sequence Number
  - Array of memory chunks providing capacity for memory region
    - Chunk Offset within memory region
    - Link to memory chunk


```
{
    "Id": "1",
    "Name": "CXL Memory Region",
    "Description": "CXL Memory Region",
    "Status": {
        "State": "Enabled",
        "Health": "OK",
        "HealthRollup": "OK"
    },
    "RegionType": "Static | Dynamic",
    "RegionNumber": 0,
    "RegionBaseOffsetMiB": 8192,
    "RegionSizeMiB": 65536,
    "ShareableRegion": false,
    "NonVolatileRegion": false,
    "HardwareManagedCoherencyRegion": true,
    "SanitizeOnRelease": false,
    "BlockSizeMiB": 128,
    "ExtentsCount": 1,
    "MemoryExtents": [
        {
            "ExtentOffsetMiB": 1024,
            "ExtentSizeMiB": 4096,
            "Tag": "User Defined Tag",
            "SequenceNumber": 0
        }
    ],
    "MemoryChunks": [
        {
            "ChunkOffsetMiB": 1024,
            "ChunkLink": {
                "@odata.id": "/redfish/v1/Chassis/CXL/MemoryDomain/1/MemoryChunks/1"
            }
        }
    ],
    "Oem": {},
    "@odata.id": "/redfish/v1/Chassis/CXL/PCIeDevices/1/CXLLogicalDevices/1/MemoryRegions/1"
}
```

PCIe Device Schema

- Existing schema updated with new property – Dynamic Capacity – containing DCD attributes
  - Number of hosts supported by the dynamic capacity device
  - Number of memory regions supported by the device dynamic capacity
  - Total dynamic memory capacity of the device
  - List of supported Add Capacity Policies
  - List of supported Remove Capacity Policies
  - List of supported Memory Block Sizes per Memory Region
  - Flag indicating support for memory sanitization on release per Memory Region


```
{
    "Id": "1",
    "Name": "CXL DCD",
    "Description": "CXL Device supporting dynamic capacity",
    "CXLDevice": {
        "DynamicCapacity":{
            "MaxHosts": 1,
            "MaxDynamicCapacityRegionsNumber": 1,
            "TotalDynamicCapacityMiB": 65536,
            "AddCapacityPoliciesSupported":[
                "Free", "Contiguous", "Prescriptive"
            ],
            "ReleaseCapacityPoliciesSupported":[
                "Tag-based", "Prescriptive"
            ],
            "MemoryBlockSizesSupported":[
                {
                    "RegionNumber": 0,
                    "BlockSizeMiB": [ 64, 128, 256, 512, 1024 ]
                }
            ],
            "SanitizationOnReleaseSupported":[
                {
                    "RegionNumber": 0,
                    "SanitizationConfigurable": true
                }
            ]
        }
    },
    "@odata.id": "/redfish/v1/Chassis/CXL/PCIeDevices/1"
}
```

CXL Logical Device Schema

- Existing schema updated with new properties
  - Link to collection of Memory Regions defined for CXL Logical Device

```
{
    "Id": "1",
    "Name": "CXL Logical Device",
    "Description": "CXL Logical Device supporting dynamic memory capacity",
    "MemoryRegions": {
        "@odata.id": "/redfish/v1/Chassis/CXL/PCIeDevices/1/CXLLogicalDevices/1/MemoryRegions"
    },
    "@odata.id": "/redfish/v1/Chassis/CXL/PCIeDevices/1/CXLLogicalDevices/1"
}
```

Memory Schema

- Existing schema updated with new properties
  - Array of links to Memory Regions providing media for logical memory device

```
{
    "Id": "1",
    "Name": "System CXL logical memory device",
    "Description": "System logical memory device representing remote CXL Memory Region",
    "Links": {
        “MemoryRegionMediaSources": [
            {
                "@odata.id": "/redfish/v1/Chassis/CXL/PCIeDevices/1/CXLLogicalDevices/1/MemoryRegions/1"
            }
        ]
    },
    "@odata.id": "/redfish/v1/Systems/1/Memory/1"
}
```

Memory Chunks Schema

- Existing schema updated with new properties
  - Array of links to Memory Regions exposing memory provided by this Memory Chunk

```
{
    "Id": "1",
    "Name": "Memory Chunk 1",
    "Description": "Physical Memory Chunk providing capcity for CXL Memory Regions",
    "Links": {
        "MemoryRegions": [
            { "@odata.id": "/redfish/v1/Chassis/CXL/PCIeDevices/1/CXLLogicalDevices/1/MemoryRegions/1" }
        ]
    },
    "@odata.id": "/redfish/v1/Chassis/CXL/MemoryDomains/1/MemoryChunks/1"
}
```

Connection Schema

- Existing schema updated with new properties
  - Memory Region Info defining connection properties for memory region


```
{
    "Id": "1",
    "Name": "Connection 1",
    "Description": "CXL Memory Region Connection",
    "ConnectionType": "Memory",
    "MemoryRegionInfo": [
        {
            "AccessCapabilities": [
                "Read",
                "Write"
            ],
            "MemoryRegion": {
                "@odata.id": "/redfish/v1/Chassis/CXL/PCIeDevices/1/CXLLogicalDevices/1/MemoryRegions/1"
            }
        }
    ],
    "@odata.id": "/redfish/v1/Fabrics/CXL/Connections/1"
}
```

Additional Resources

- More details of the CXL model can be found at the CXL Public Mockup on the DMTF Website 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

Thank you!

For more information,
visit us online at
dmtf.org

Visit the
Redfish Developer Hub at
redfish.dmtf.org