



OPEN

Compute Project

PCIe Extended Connectivity Requirements Document

Revision 0.15
May 25th, 2023

Contributing Company – Contact Person

Seagate Technologies – Mohamad El-Batal

Ayar-Labs – LK Bhupathi

Astera-Labs – Richard Ward

Revision History

Version	Date	Details
0.1	1/3/2023	First draft – Mohamad El-Batal
0.11	1/9/2023	Added Top level rack diagram and Out-Of-Band Security/Management
0.12	2/12/2023	Added channel model diagrams for DAC and AEC
0.13	2/13/2023	Added Rack level Diagrams for CXL Scenarios
0.14	2/23/2023	Added cabling suggestions for Sideband signals & changed title from External Connectivity to Extended Connectivity

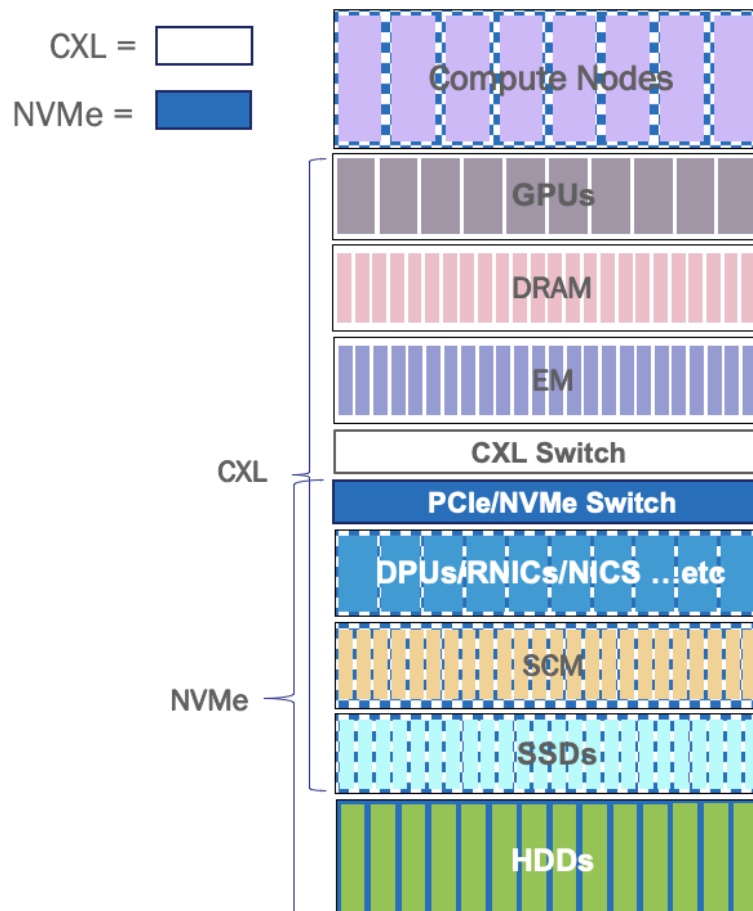
Table of Contents

1 OCP	4
1.1 Scope	4
1.2 ACKNOWLEDGEMENT	5
2 COMPLIANCE W/ OCP TENETS	6
2.1 Openness	6
2.2 Efficiency	6
2.3 Impact	6
2.4 Scale	6
3 REFERENCES	7
4 External PCIe Connectivity Scenarios for CXL	8
4.1 Compute Cluster	15
4.2 Compute to Accelerator	15
4.3 Accelerator Cluster	15
4.4 Compute to Memory	15
4.5 Accelerator to Memory	15
4.6 Memory Cluster	16
5 External PCIe Connectivity Scenarios for NVMe	16
5.1 Compute to Storage	16
5.2 Accelerator to Storage	17
5.3 Storage Cluster	17
6 External Connectivity Types	19
6.1 Direct Attached Connection (DAC)	19
6.2 ACTIVE ELECTRICAL CONNECTION (AEC)	21
6.3 Direct Optical Connection (DOC)	22
6.4 Active Optical Connection (AOC)	23
6.5 Activity Signal	23
6.6 SYSTEM DESIGN	25
7 Power	25
7.1 Connector Power/Colling Limits	25
8 SECURITY	25
8.1 Security Requirements from OCP Datacenter NVMe SSD Specification	25

1 OCP

1.1 Scope

To document the various industry usage-model scenarios and requirements for disaggregated NVMe and CXL inter-connected devices including: Switch, Compute, Accelerator, Networking, Memory, and Storage modules using PCIe Gen5 & Gen6 external Direct Attached Connection (DAC), Active Electrical Connection (AEC), Direct Optical Connection (DOC), and Active Optical Connection (AOC) Cables. We shall not explore and/or specify any specific detailed connector or cabling implementations specification at this time; however, a future industry standards committee will be formed by a team of willing members aiming to materialize the requirement output of this Workstream into a complete standardized cabling and connector specification that leverages existing connector and cabling designs and provide the most optimal solution level Total Cost of Ownership (TCO) results and the best industry wide economy of scale benefits. This workstream shall explore and document only the most relevant external PCIe physical layer connectivity scenarios that are currently in plan or are foreseen for future datacenter and enterprise applications in conjunction with NVMe and CXL protocol based external connectivity applications that are intra-rack and/or inter-rack in scale.



1.2 Acknowledgement

The OCP External PCIe Connectivity Requirements Document was created with collaboration from many OCP member companies and facilitated by the OCP External PCIe Connectivity Workstream under the OCP Server Project Workgroup. The OCP External PCIe Connectivity Requirements Workstream would like to acknowledge the following member companies for their contributions to the External PCIe Connectivity Requirements Document:

OCP Member Company
Amphenol Corporation
IIVI/Coherent Corporation
TE Connectivity Corporation
Molex Connectivity Corporation
Samtec Connectivity corporation
Broadcom Limited
Marvell Semiconductor, Inc.
Google
Meta
Microsoft
Molex Corporation
Quanta
Seagate Technology Corporation
Super Micro Corporation
Western Digital Corporation
Microchip Technology
Astera-Labs Corporation
Serial Cables Corporation
Ayar Labs Corporation

2 Compliance w/ OCP Tenets

2.1 Openness

The OCP External PCIe Connectivity Requirements Document was developed in dialogue with the entire OCP Server, Datacenter, Memory and Storage Project membership and the resulting requirements document represents the initial requirement set for this new category of external PCIe physical layer connectivity solutions. This ecosystem is nascent, so we are confident that this document will evolve over time, and that this requirements document represents the wisdom of the collective ecosystem on the scenarios and requirements for creating NVMe and/or CXL based solutions over external PCIe physical layer connectivity in the near future datacenter and enterprise applications.

2.2 Efficiency

The efficiency goals of this workstream are to identify the key scenarios and requirements that connector and cabling manufacturers across the industry can build external PCIe physical layer connectivity solutions that follow a consistent and unified set of requirements, vetted by vendors and customers. Architecturally, such external PCIe connectivity infrastructure enables datacenter and enterprise solution designers to take advantage of using the same connectivity infrastructure and drive the best economy of scale TCO benefits to the NVMe and CXL connectivity industry. Additionally, it is not a preset requirement to define unified connectivity requirements for both NVMe and CXL, but to explore both protocol solution scenarios separately and then explore any potential intersecting solution requirements along with the pros and cons of such possible unified solutions if any. Solution level Total Cost of Ownership (TCO) optimization is king, and any decision to unify requirements must pass the TCO analysis checkpoint.

2.3 Impact

The standardization of External PCIe physical layer connectivity for NVMe and CXL will have a huge impact on the overall datacenter and enterprise ecosystems. It shall extend the once considered internal server centric PCIe converged bus to enabling the PCIe physical layer to become the ideal high-bandwidth and ultra-low-Latency optimized disaggregation interface of choice for future NVMe and CXL protocol interconnected compute, acceleration, memory, and storage solutions.

2.4 Scale

The scale of this new interface connectivity ecosystem will be governed by the PCIe physical layer attributes at Gen5 and Gen6 differently; however, the scale limits will be different relative to each of the two protocols.

3 References

From <https://www.pcisig.com>:

- PCI Express® Base Specification, Revision 6.0
- PCI Express® Base Specification, Revision 5.0
- ECN - Component Measurement and Authentication (CMA) (June 25, 2020)

From <https://nvmexpress.org/>:

- NVM Express® Base Specification revision 2.0b
- NVM Express Management Interface Specification, Revision 1.2b
- NVM Express NVM Command Set Specification, Revision 1.0b
- NVM Express Zoned Namespace Command Set Specification, Revision 1.1b
- NVM Express Key Value Command Set Specification, Revision 1.0b
- NVM Express NVMe over PCIe Transport Specification, Revision 1.0b

From <https://nvmexpress.org/>:

- NVM Express® Base Specification revision 2.0b
- NVM Express Management Interface Specification, Revision 1.2b
- NVM Express NVM Command Set Specification, Revision 1.0b
- NVM Express Zoned Namespace Command Set Specification, Revision 1.1b
- NVM Express Key Value Command Set Specification, Revision 1.0b
- NVM Express NVMe over PCIe Transport Specification, Revision 1.0b

From <https://www.opencompute.org/>:

- Datacenter NVMe® SSD Specification, Revision 2.0
- Storage NVMe® HDD Specification, Revision 0.7

From <https://trustedcomputinggroup.org/>:

- TCG Storage Security Subsystem Class: Opal, Version 2.02, Revision 1.0
- TCG Storage Interface Interactions Specification (SIIS), Version 1.11
- TCG DICE Attestation Architecture, Version 1.00

From <https://dmtf.org>

- DSP0238 Management Component Transport Protocol (MCTP) PCIe VDM Transport Binding Specification
- DSP0274 Security Protocol and Data Model (SPDM) Specification, Version 1.2.0

4 External PCIe Connectivity Scenarios for CXL

4.1 U-Heights of Various CXL Related Chassis

- 1U Servers and Switches
- 2U Servers, GPUs blades, EDSFF Memory
- 3U PCIe CEM cards, NICs, RNICs, DPUs, iPUs ...etc.
- 4U Enterprise Servers
- 5U Potential 19" Half-width chassis vertically mounted solutions
- 6U Potential 21" Half-width chassis vertically mounted solutions

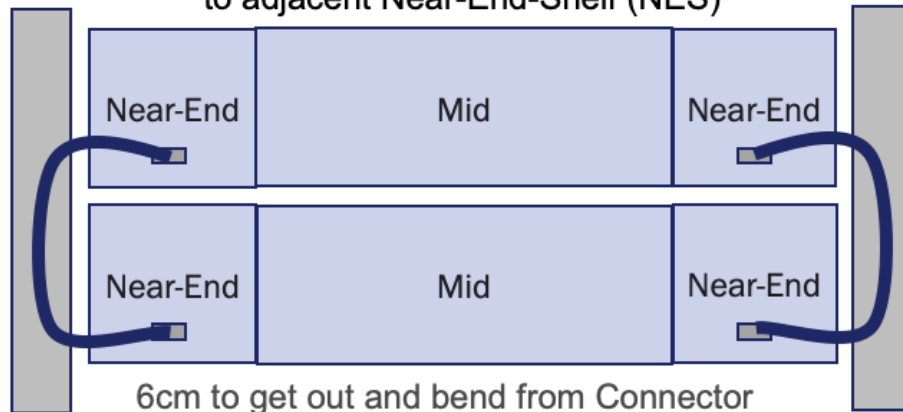
4.2 Rack Level Configuration Considered

The cable shall come in different widths and lengths to allow for various connectivity scenarios

0.5M from Near-End-Shelf (NES) to adjacent Near-End-Shelf (NES)

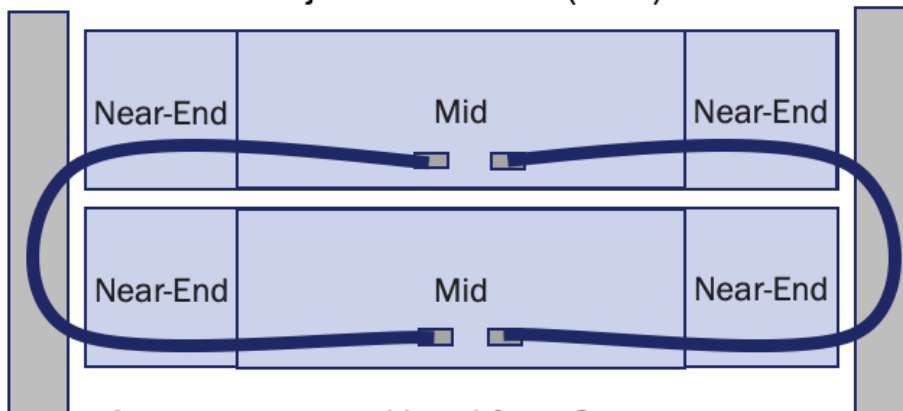
1.0M from Mid-Shelf (MS) to adjacent Mid-Shelf (MS) (w/o CMA)

**~0.5M cable from Near-End-Shelf (NES)
to adjacent Near-End-Shelf (NES)**



6cm to get out and bend from Connector
 10cm to go to the rack rails and bend
 18cm between 2x 4U shelves
 10cm to go to the rack rails and bend
 6cm to get bend and get into Connector

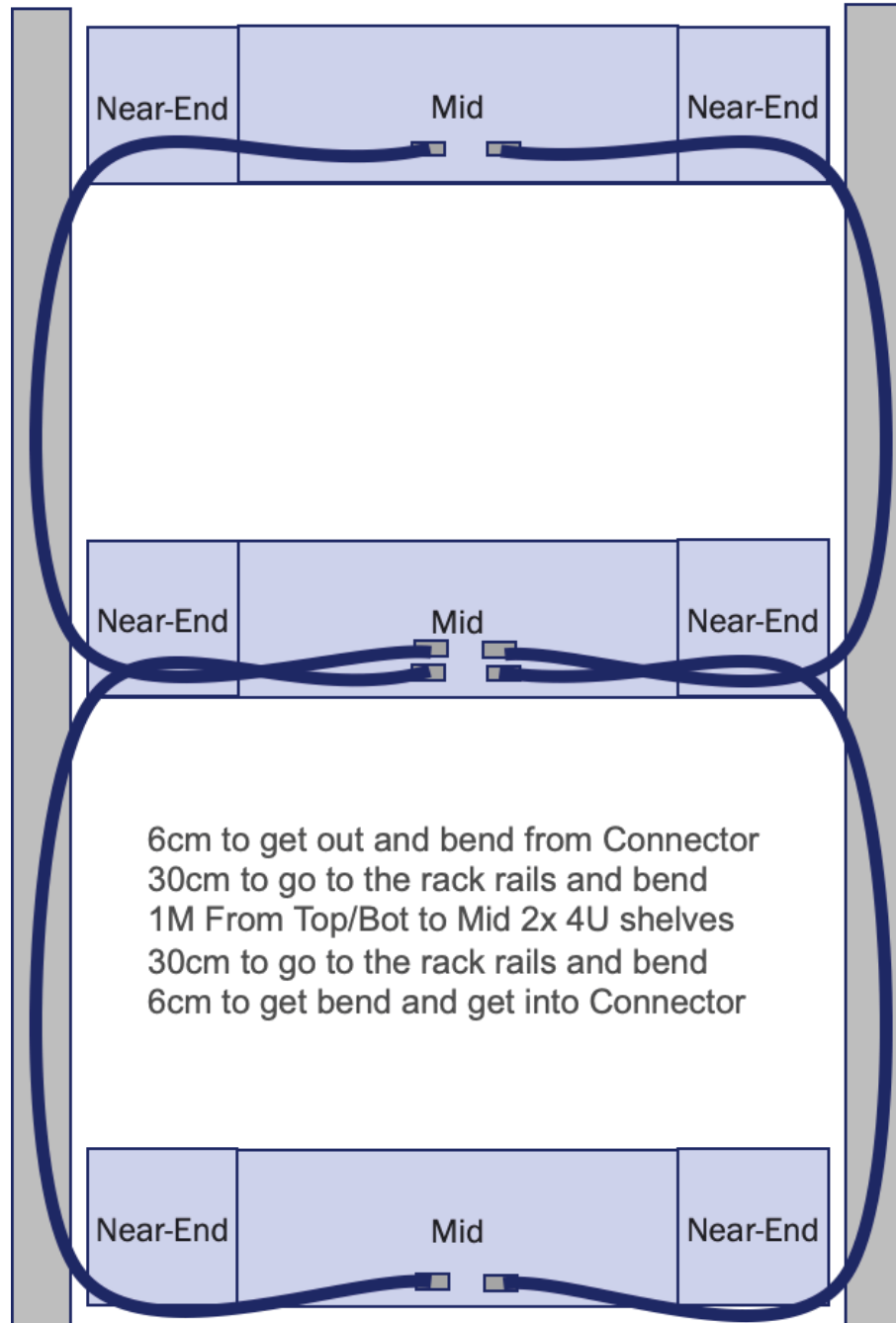
**~1.0M cable from Mid-Shelf (MES)
to adjacent Mid-Shelf (MES)**



6cm to get out and bend from Connector
 30cm to go to the rack rails and bend
 18cm between 2x 4U shelves
 30cm to go to the rack rails and bend
 6cm to get bend and get into Connector

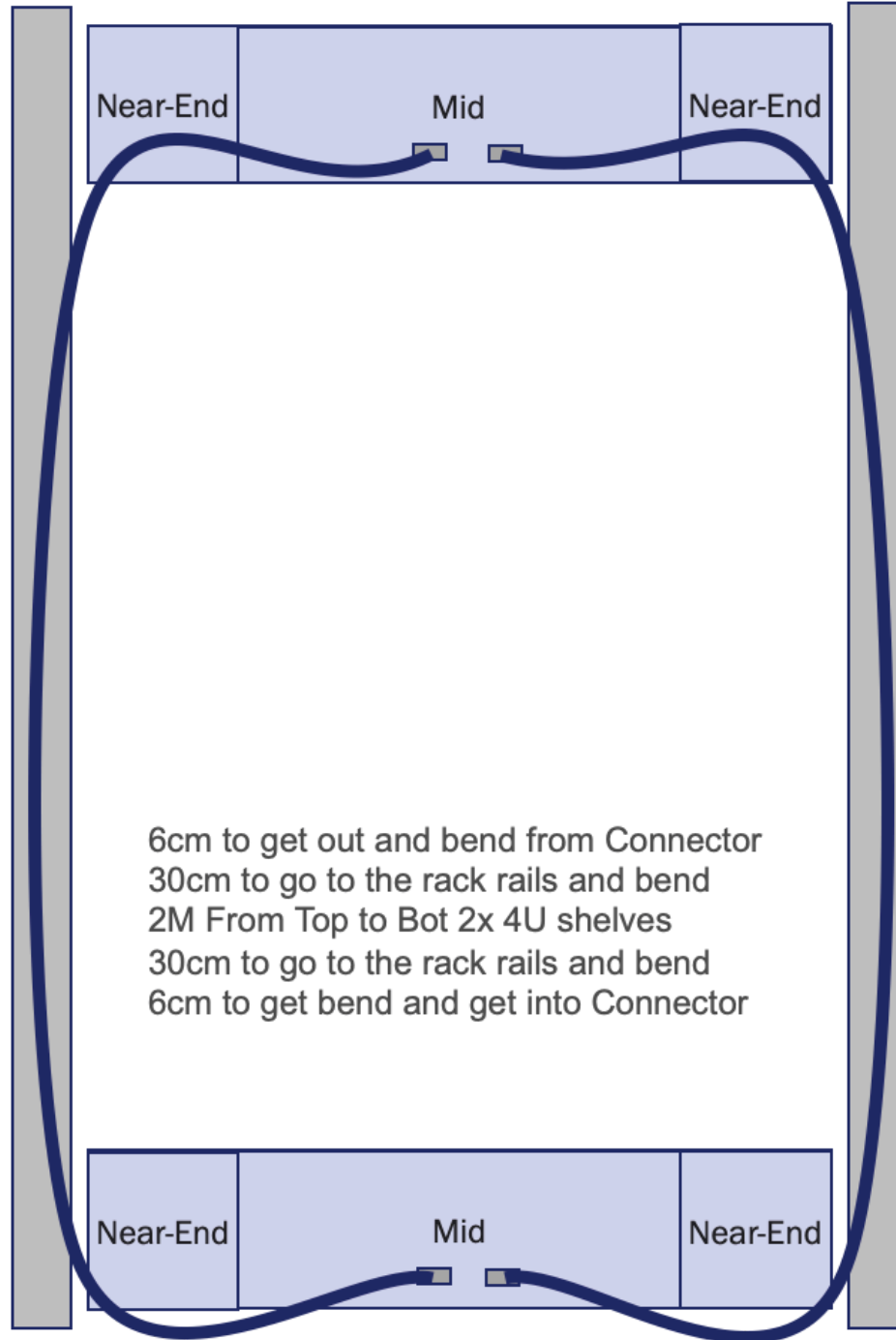
2.0M from Mid-Of-Rack (MOR) to Top/Bott-Of-Rack (T/BOR)

~2.0M cable from Mid-Of-Rack (MOR) to Top/Bott-Of-Rack (T/BOR)

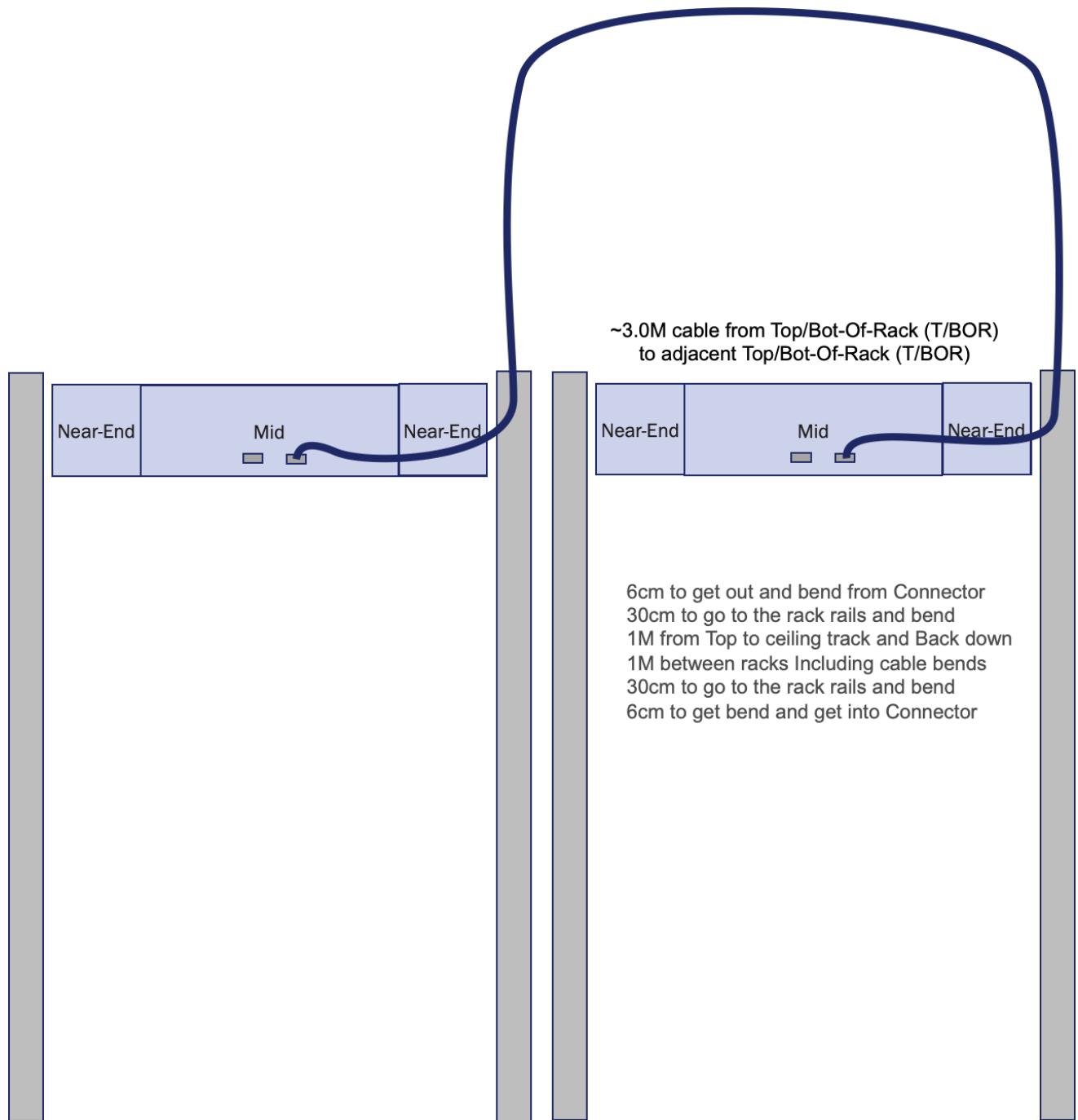


2.0M from Bott/Top-Of-Rack (B/TOR) to Top/Bott-Of-Rack (T/BOR)

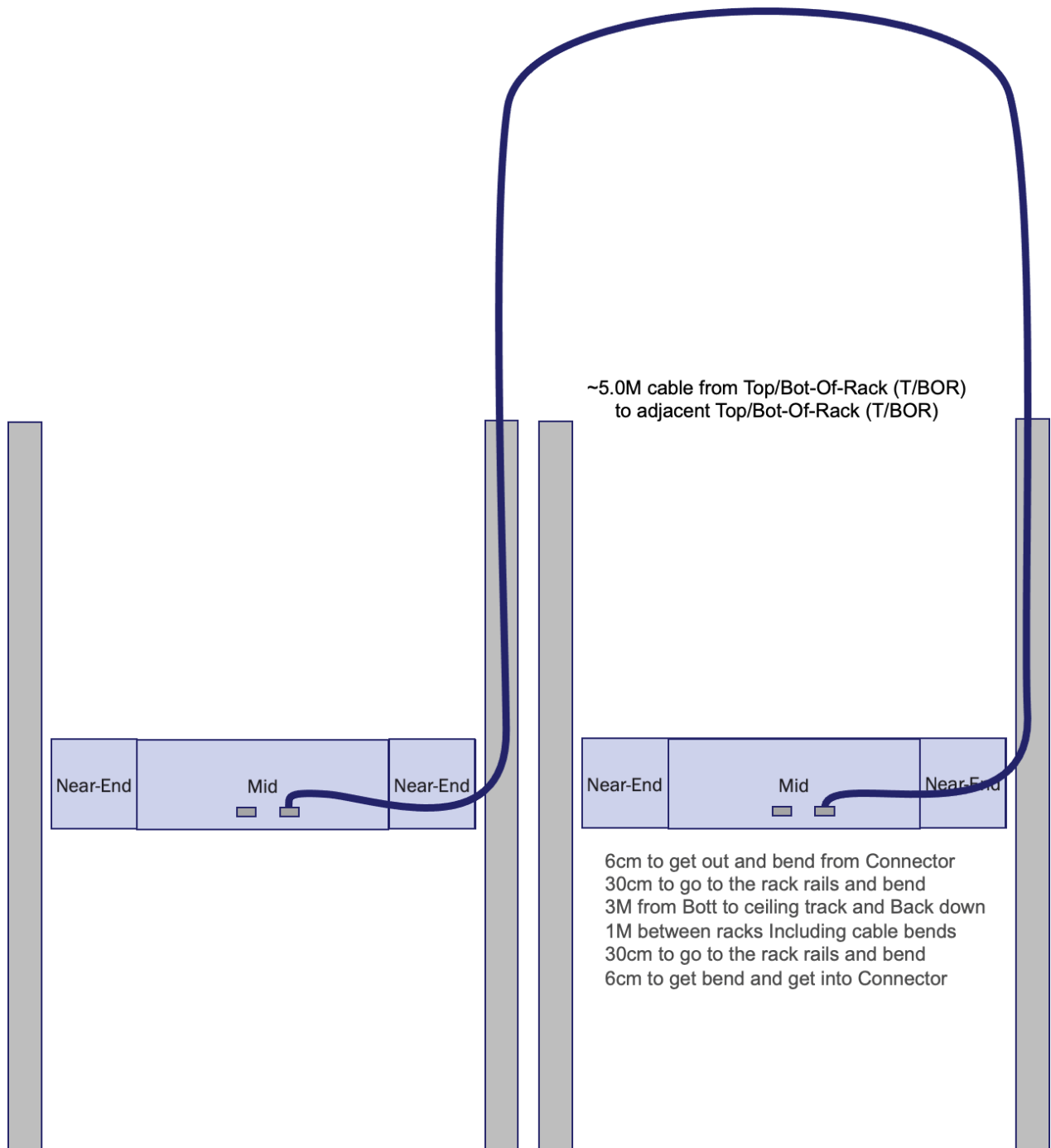
~3.0M cable from Mid-Of-Rack (MOR)
to Top/Bott-Of-Rack (T/BOR)



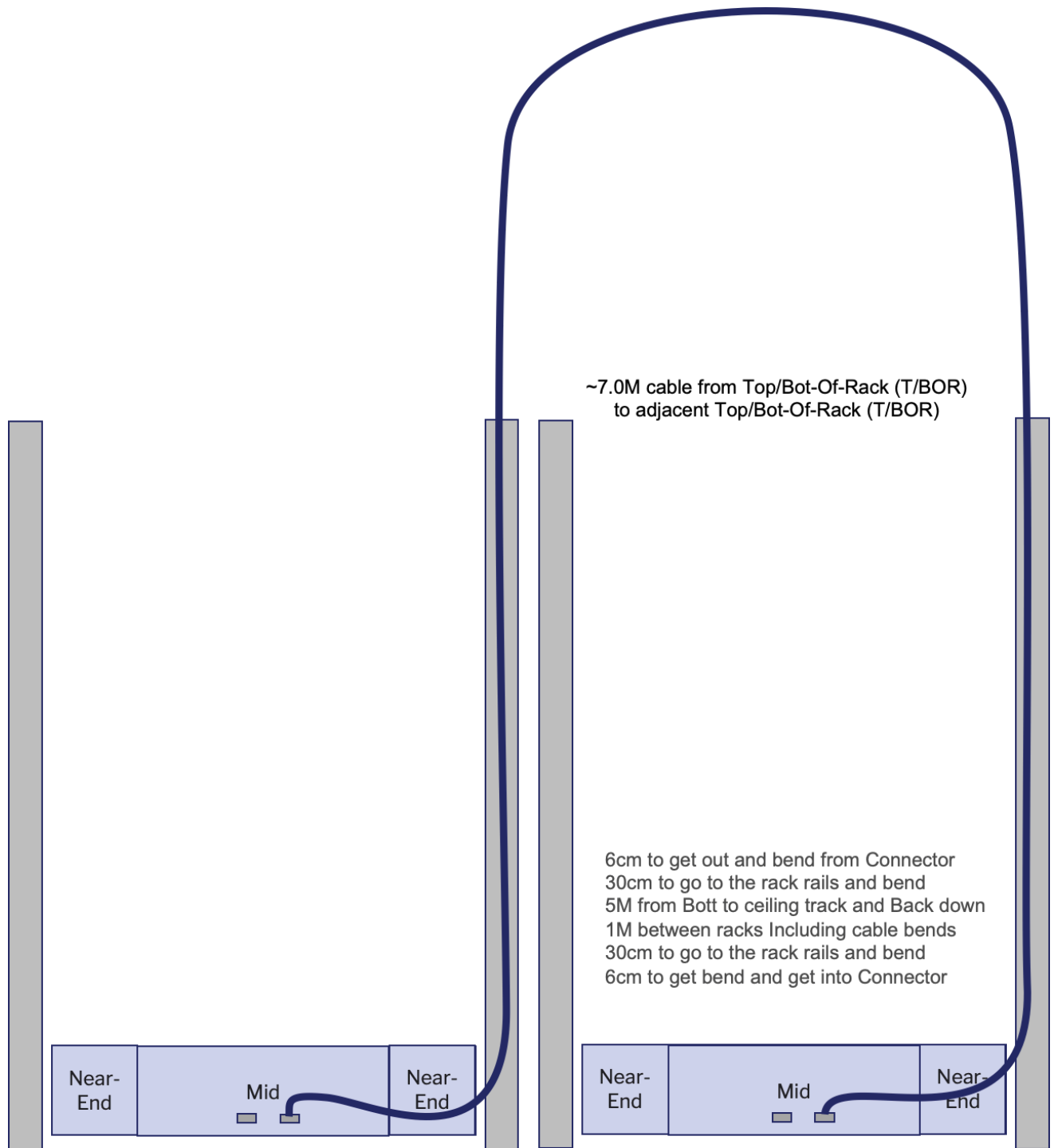
3.0M from Top-Of-Rack (TOR) to adjacent Top-Of-Rack (TOR) w/ Ceiling Tracks
3.0M from Bott-Of-Rack (BOR) to adjacent Bott-Of-Rack (BOR) w/ Subfloor Tracks



3.0M from Mid-Of-Rack (MOR) to adjacent Mid-Of-Rack (MOR) w/ Ceiling of Subfloor Tracks



5.0M from Top-Of-Rack (TOR) to adjacent Bott-Of-Rack (BOR) w/ Ceiling of Subfloor Tracks
 5.0M from Bott-Of-Rack (BOR) to adjacent Top-Of-Rack (TOR) w/ Ceiling of Subfloor Tracks



7.0M from Top-Of-Rack (TOR) to adjacent Top-Of-Rack (TOR) w/ Subfloor Tracks
 7.0M from Bott-Of-Rack (BOR) to adjacent Bot-Of-Rack (TOR) w/ Ceiling Tracks

4.3 Compute Cluster

Req ID	Speed Gen	Config/Width	Lengths
Scenario-1	PCIe-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-2	PCIe-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M

4.4 Compute to Accelerator

Req ID	Speed Gen	Config/Width	Lengths
Scenario-3	PCIe-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-4	PCIe-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M

4.5 Accelerator Cluster

Req ID	Speed Gen	Config/Width	Lengths
Scenario-4	PCIe-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-5	PCIe-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M

4.6 Compute to Memory

Req ID	Speed Gen	Config/Width	Lengths
Scenario-6	PCIe-Gen5	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-7	PCIe-Gen5	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-8	PCIe-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-9	PCIe-Gen6	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-10	PCIe-Gen6	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-11	PCIe-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M

4.7 Accelerator to Memory

Req ID	Speed Gen	Config/Width	Lengths
Scenario-12	PCIe-Gen5	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-13	PCIe-Gen5	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-14	PCIe-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-15	PCIe-Gen6	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-16	PCIe-Gen6	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-17	PCIe-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M

4.8 Memory Cluster

Req ID	Speed Gen	Config/Width	Lengths
Scenario-18	PCIe-Gen5	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-19	PCIe-Gen5	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-20	PCIe-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-21	PCIe-Gen6	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-22	PCIe-Gen6	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M
Scenario-23	PCIe-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M & 7.0M

5 External PCIe Connectivity Scenarios for NVMe

The cable shall come in different widths and lengths to allow for various connectivity scenarios

- 0.5M from Near-End-Shelf (NES) (w/o CMA) to adjacent Near-End-Shelf (NES) (w/o CMA)
- 1.0M from Far-End-Shelf (FES) (w/o CMA) to adjacent Far-End-Shelf (FES) (w/o CMA)
- 2.0M from Near-End-Rack (TOR) (w/o CMA) to adjacent Near-End-Shelf (NES) (w/ CMA)
- 2.0M from Mid-Of-Rack (MOR) (w/o CMA) to Top/Bott-Of-Rack (T/BOR) (w/o CMA)
- 3.0M from Bott-Of-Rack (BOR) (w/o CMA) to Top-Of-Rack (TOR) (w/o CMA)
- 3.0M from Top-Of-Rack (TOR) (w/o CMA) to adjacent Top-Of-Rack (TOR) (w/o CMA)
- 5.0M from Top-Of-Rack (TOR) (w/o CMA) to adjacent Top-Of-Rack (TOR) (w CMA)
- 5.0M from Mid-Of-Rack (MOR) (w/o CMA) to adjacent Mid-Of-Rack (MOR) (w/o CMA)
- 7.0M from Bott-Of-Rack (BOR) (w/o CMA) to adjacent Bott-Of-Rack (BOR) (w/o CMA)
- 7.0M from Mid-Of-Rack (MOR) (w/o CMA) to adjacent Mid-Of-Rack (MOR) (w CMA)
- 10.0M from Bott-Of-Rack (BOR) (w/o CMA) to adjacent Bott-Of-Rack (BOR) (w/ CMA)
- 10.0M from Bott-Of-Rack (BOR) (w/ CMA) to adjacent Bott-Of-Rack (BOR) (w/ CMA)

5.1 Compute to Storage

Req ID	Speed Gen	Config/Width	Lengths
Scenario-24	PCIe-Gen5	1x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-25	PCIe-Gen5	2x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-26	PCIe-Gen5	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-27	PCIe-Gen5	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-28	PCIe-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-29	PCIe-Gen6	1x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-30	PCIe-Gen6	2x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-31	PCIe-Gen6	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-32	PCIe-Gen6	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-33	PCIe-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M

5.2 Accelerator to Storage

Req ID	Speed Gen	Config/Width	Lengths
Scenario-34	PCIe-Gen5	1x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-35	PCIe-Gen5	2x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M

Scenario-36	PCle-Gen5	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-37	PCle-Gen5	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-38	PCle-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-39	PCle-Gen6	1x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-40	PCle-Gen6	2x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-41	PCle-Gen6	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-42	PCle-Gen6	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-43	PCle-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M

5.3 Storage Cluster

Req ID	Speed Gen	Config/Width	Lengths
Scenario-44	PCle-Gen5	1x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-45	PCle-Gen5	2x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-46	PCle-Gen5	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-47	PCle-Gen5	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-48	PCle-Gen5	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-49	PCle-Gen6	1x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-50	PCle-Gen6	2x4	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-51	PCle-Gen6	1x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-52	PCle-Gen6	2x8	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M
Scenario-53	PCle-Gen6	1x16	0.5M, 1.0M, 2.0M, 3.0M, 5.0M, 7.0M & 10.0M

6 External Connectivity Types

6.1 Direct Attached Connection (DAC)

DAC cables must be completely passive on data-path signals but must contain an SMBus target controller and signals to provide Vendor Product Data (VPD) information on both ends.

Out-of-Band management between devices on extended connectivity solutions shall not use Sideband signals like I2C or I3C, which add more wires across the cable thus more overall cost.

Such Out-of-Band (control-plane) management communication could make use of standard PCIe out-of-band transports that are available to us from various Board Management Controllers (BMCs) devices with Management Component Transport Protocol (MCTP) support. Firmware downloads and Reset functions can also be done using MCTP over DOE or over VDM transports commands. Security attestation could also make use of Security Protocol and Data model (SPDM) over PCIe Data Object Exchange (DOE) or over MCTP over Vendor Defined Messages (VDM) transport layer.

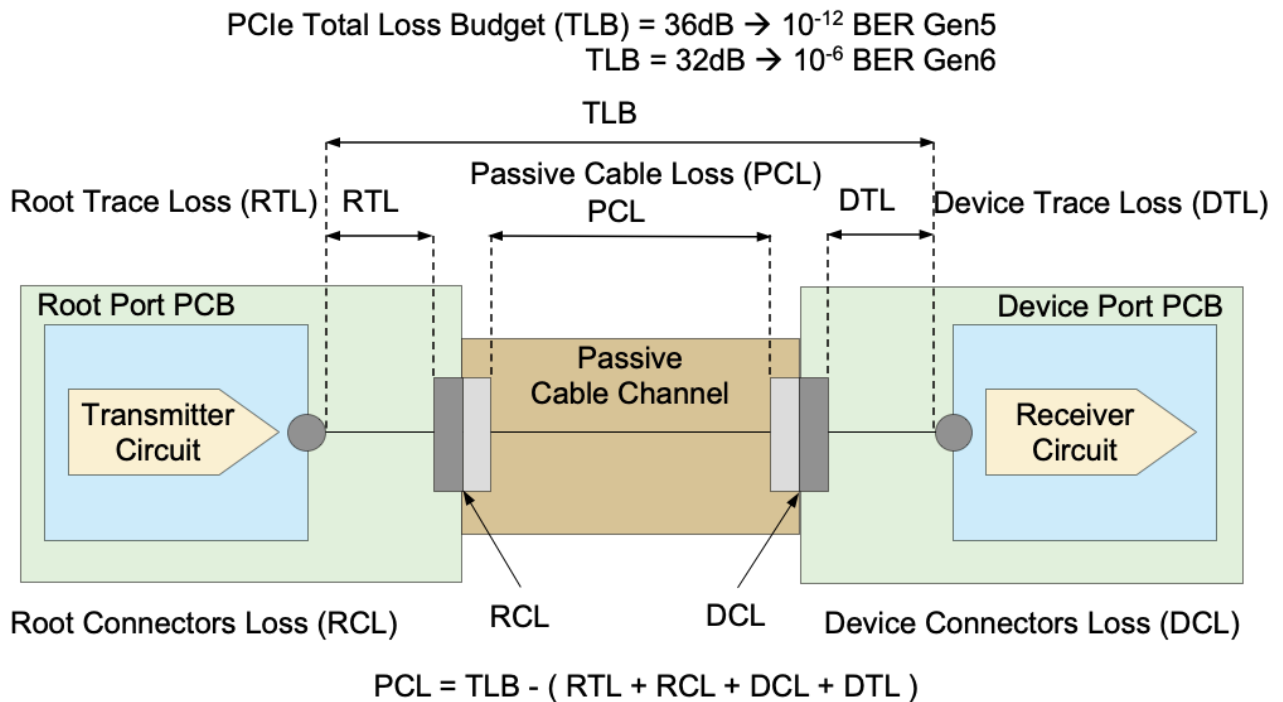
Alternatively, users can make use of other sideband cabled interfaces between connected devices over Ethernet or USB to enable Security, Out-of-band Management, Firmware download or Reset functions.

DAC cables shall require host and endpoint support for SRIS with Spread Spectrum Clocking (SSC) and SRNS with standard single frequency clocking both as mandatory functions in order to save on cabling cost and EMI reduction.

Alternatively, a shared synchronization clock can be connected between Root and Endpoint devices over the same sideband control-plane management interface or a separate clock synchronization cable if desired by the architecture

DAC cables must allow for standard PCIe physical layer signal requirement including connector discontinuities and both PCB routes on both the Source and Device ends.

DAC cables shall support PCIe port bifurcation but in SRIS mode only and without the additional PCIe differential REFCLK(+/-) and PERST# signals.



The Printed Circuit Board (PCB) dB budget loss depends greatly on the dielectric material used, copper trace width and weight, and the system designer must tune their solution specification to achieve the required balance in the above equation not to exceed the 36dB budget of the specific PCIe generation targeted.

There is more to signal integrity than the loss budget, since many factors play a big role in how long of a cable can be used including the trace routing and impedance matching as well as any reflections or crosstalk. The system designer must qualify the complete end to end solution.

The Passive Cable Channel dB loss budget also depends greatly on the Twinax cable gauge and type used. The DAC cable provider must specify their loss per distance as well as their connector's loss.

Cable AWG	IL* (db/m) @16GHz (Gen5 NRZ & Gen6 PAM4)
28	4.3
30	5.3
32	7.2

DAC cables can use any Micro-Twinax cabling that matches the connector and PCBs impedances on both ends. Depending on the cable characteristics used the system designer must the total distance and maintain the PCL requirements dictated by the system designer.

DAC cables can optionally support different connector width on each end to allow for Octopus cable for 1 x8 to 2 x4, and 1 x16 to 2 x8 or 4 x4 configurations with PCIe bifurcated connections.

Note: Add the different cable bifurcation scenarios

Req ID	DAC	#P	RT	SMBus	SRIS/NS	CuCab	CDmm	Lnths M	BndRad
CNNCT-1	1x4	1	No	Yes	Optional	30-AWG	1x a?	Up to 1	5x CD
CNNCT-2	1x8	1	No	Yes	Optional	28-AWG	1x b?	Up to 1	5x CD
CNNCT-3	1x16	1	No	Yes	Optional	28-AWG	2x b?	Up to 1	5x CD
CNNCT-4	x8(2x4)	2	No	Yes	Optional	28-AWG	1x b?	Up to 1	5x CD
CNNCT-5	x16(2x8)	2	No	Yes	Optional	28-AWG	2x b?	Up to 1	5x CD
CNNCT-6	x16(4x4)	4	No	Yes	Optional	28-AWG	2x b?	Up to 1	5x CD
CNNCT-7	x8 to 2x4	2	No	Yes	Optional	28-AWG	2x a?	Up to 1	5x CD
CNNCT-8	x16 to 2x8	2	No	Yes	Optional	28-AWG	2x b?	Up to 1	5x CD
CNNCT-9	x16 to 4x4	4	No	Yes	Optional	28-AWG	4x a?	Up to 1	5x CD

6.2 Active Electrical Connection (AEC)

AEC cables support Re-Timer devices on both end modules, forming an Active-Active connection type.



Each Re-Timer that is added into the path resets the PCIe link budget, allowing for an extra 36dB channel loss (PCIe-Gen5 for 1×10^{-12} BER) or 32dB channel loss (PCIe-Gen6 for 1×10^{-6} BER) in the system, with equivalent reach extension.

Out-of-Band management between devices on extended connectivity solutions shall not use sideband signals such as I2C, which add more wires across the cable thus more overall cost and is not scalable for the optical solutions

Such Out-of-Band (control-plane) management communication could make use of standard PCIe out-of-band transports that are available to us from various Board Management Controllers (BMCs) devices with Management Component Transport Protocol (MCTP) support. Firmware downloads and Reset functions can also be done using MCTP over DOE or over VDM transports commands. Security attestation could also make use of Security Protocol and Data model (SPDM) over PCIe Data Object Exchange (DOE) or over MCTP over Vendor Defined Messages (VDM) transport layer.

Alternatively, users can make use of other sideband cabled interfaces between connected devices over Ethernet or USB to enable Security, Out-of-band Management, Firmware download or Reset functions.

AEC cables shall require host and endpoint support for SRIS with Spread Spectrum Clocking (SSC) and SRNS with standard single frequency clocking both as mandatory functions in order to save on cabling cost and EMI reduction.

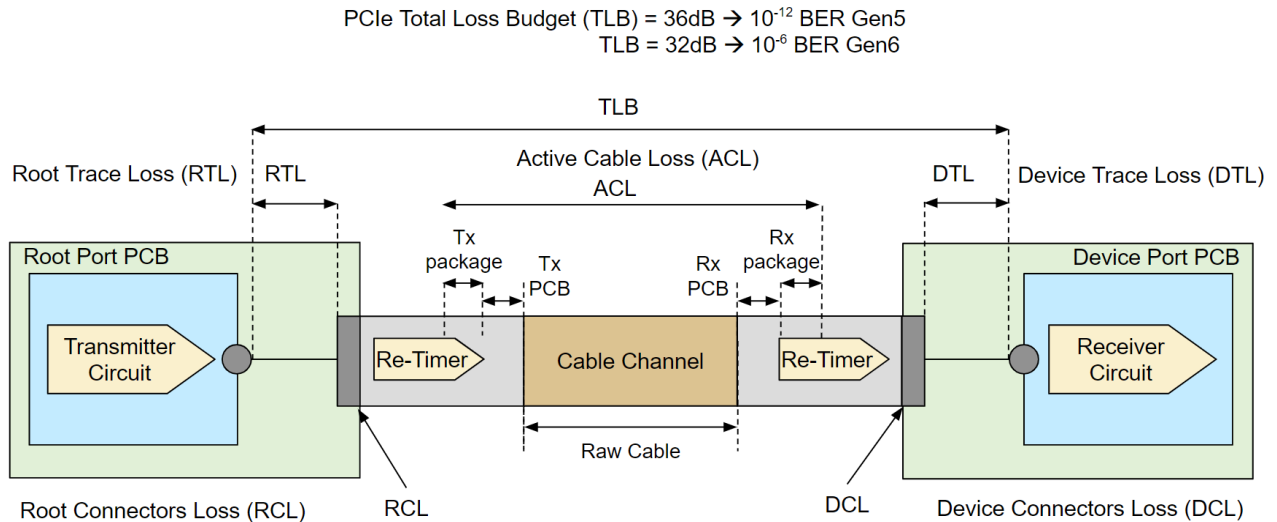
Alternatively, a shared synchronization clock can be connected between Root and Endpoint devices over the same sideband control-plane management interface or on a separate clock synchronization cable if desired by the architecture

AEC cables must allow for standard PCIe physical layer signal requirement including connector discontinuities and both PCB routes on both the Source and Device ends.

AEC cables shall support PCIe port bifurcation but in SRIS mode only and without the additional PCIe differential REFCLK(+/-) and PERST# signals.

AEC cables must contain an SMBus target controller and signals to provide Vendor Product Data (VPD) information on both ends but not passed over the cable between the two connectors. It is

proposed the PCIe module control follows CMIS functionality for the module, otherwise we'll need to define our own set of register pages and functions (for example for FW upgrade) into the OCP specifications.



With the additional link budget from including retimers, the maximum length of the cable will depend on the line-side Package and PCB loss and the gauge of the cable (AWG).

The Active Cable Channel dB-loss budget also depends on the Twinax cable gauge and type used. The AEC cable provider must specify their loss per distance based on their Re-Timer ability to mask the Passive Cable Loss from the Root and Device ports, but they still must provide their connector's losses after the Re-timer in order for the system designer to calculate the remaining RTL and DTL loss budgets.

AEC cables must interface and match connector impedances on both ends.

AEC cables must have the same connector type and widths on both ends of the cable.

Req ID	AEC	#P	RT	SMBus	SRIS/NS	CDmm	Lnths M	BndRad
CNNCT-10	1x8	1	Yes	Yes	Optional	1x 6.8	Up to 3	5x CD
CNNCT-11	1x16	1	Yes	Yes	Optional	2x 6.8	Up to 3	5x CD
CNNCT-12	x8(2x4)	2	Yes	Yes	Optional	1x 6.8	Up to 3	5x CD
CNNCT-13	x16(2x8)	2	Yes	Yes	Optional	2x 6.8	Up to 3	5x CD
CNNCT-14	x16(4x4)	4	Yes	Yes	Optional	2x 6.8	Up to 3	5x CD

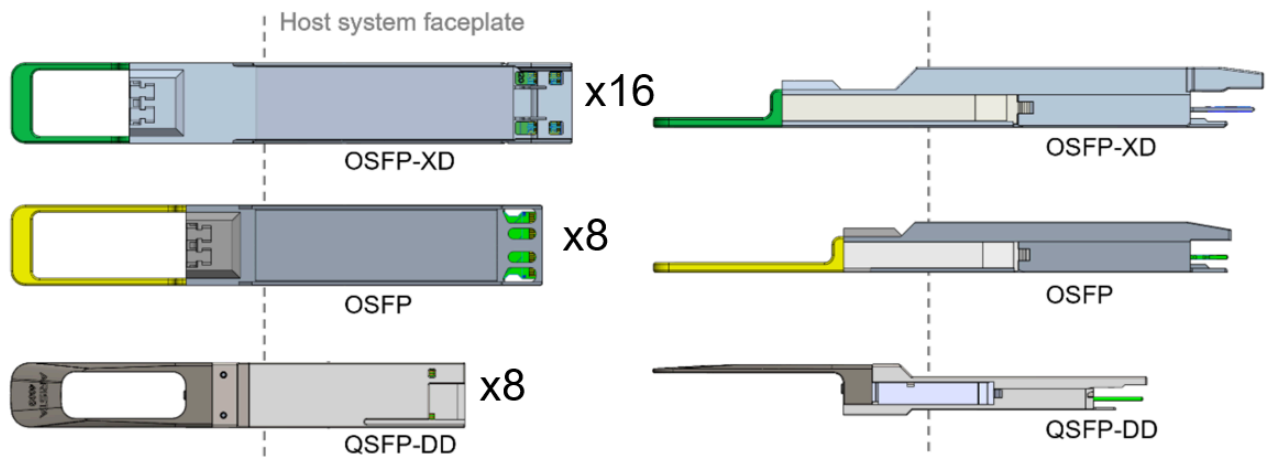
The standard (Intel defined) package form factors for PCIe-Gen5 ReTimers are:

Rate	Lanes	Package Size (mm)
Gen5	1 x 8	8.5 x 13.4
Gen5	1 x 16	8.9 x 22.8
Gen6	1 x 8	tbd
Gen6	1 x 16	tbd

These package sizes enable the Gen5 x8 solution to fit inside a QSFP-DD module case, with an expected power consumption of <12W.

Gen5 x16 would require the OSFP-XD for the electrical connectivity and the standard package size requires a 90 degree rotation to fit inside the OSFP-XD form factor, or the use of non-standard packaging.

Module form factors under consideration for the PCIe/CXL AECs are QSFP-DD, OSFP and OSFP-XD.



6.3 Direct Optical Connection (DOC)

DOC cables must be completely passive on datapath signals, and for cost optimization reasons we don't require SMBus target controller and signals to provide Vendor Product Data (VPD) information on both ends. In this scenario, optical transceiver functionality is expected to be implemented either on the system PCB or inside the package of the system SOC.

DOC cables must use Multi-Mode (MM) or Single Mode (SM) fibers.

DOC cables must support bifurcation and shall only **SRIS with Spread Spectrum Clocking (SSC) and SRNS with standard single frequency clocking.**



In typical DOC implementations, the light source that contains lasers that power the optical engine are expected to be external to the datapath devices. They could be incorporated into the system either as a module mounted on the PCB or as a pluggable like the ELSF (External Laser Small Form Factor Pluggable) module specified by OIF. Solutions maybe realized with integrated lasers if power and cost are within requirements.

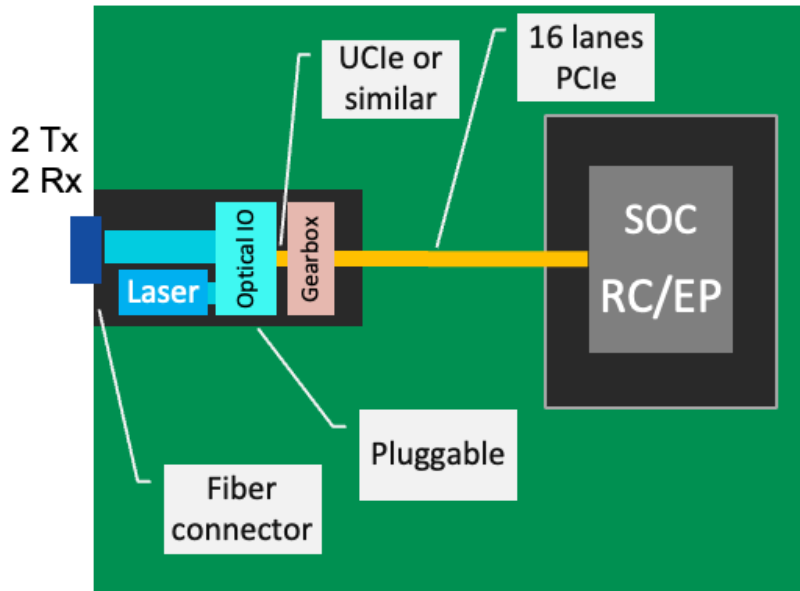
DOC cables must support different connector width on each end to allow for Octopus cable for a 1 x16 to 2 x8 configuration with PCIe bifurcated connections.

Req ID	DOC	#Ports	SMBus	SRIS/NS	OptCab	CDmm	Lnths M	BndRad
CNNCT-15	1x8	1	No	Yes	MM/SM	1x?	Up to 10	??
CNNCT-16	1x16	1	No	Yes	MM/SM	1x?	Up to 10	??
CNNCT-17	x16(2x8)	2	No	Yes	MM/SM	2x?	Up to 10	??
CNNCT-18	x16 to 2x8	2	No	Yes	MM/SM	2x?	Up to 10	??

6.4 Active Optical Connection (AOC)

AOC cables must support built-in Copper to Optical transceiver devices on both end connectors, in order to provide support for the complete standard PCIe-Gen5 at 1x10⁻¹² BER requirements, as well as PCIe-Gen6 at 1x10⁻⁶ BER requirements as far as the Physical transport layer is concerned, while keeping in mind the need for the overall connectivity solution to enhance the BER at the Link layer for the specific protocols that we are addressing being NVMe and CXL.

(x4, x8, and x16 pluggable options. Mention examples of modules that could be leveraged - QSFP, QSFP-DD and OSFP-XD. Retimer and PCIe spec reference... Power requirement? Module dimension? Gen5 could be 0.5W per retimer)



- PCIe-Gen5 BER is

AOC cable source and destination PCB route and db-loss budget must also be specified to meet the standard PCIe driver compliance requirements.

AOC cables must contain an SMBus target controller and signals to provide Vendor Product Data (VPD) information on both ends.

AOC cables must use Multi-Mode (MM) or Single-Mode (SM) Micro-Fiber cabling that comply with the optical transceiver's requirements used on both connector ends.

AOC cables must support PCIe port bifurcation in conjunction with support for SRIS with Spread Spectrum Clocking (SSC) and SRNS with standard single frequency clocking are both mandatory.

AOC cables must have the same connector type and widths on both ends of the cable.

Req ID	AOC	#P	MPO	SMBus	SRIS/NS	OptCab	CDmm	Lnths M	BndRad
CNNCT-19	1x8	1	Yes	Yes	Yes	MM/SM	1x?	5, 7 & 10	10x FD
CNNCT-20	1x16	1	Yes	Yes	Yes	MM/SM	1x?	5, 7 & 10	10x FD
CNNCT-21	x16(2x8)	2	Yes	Yes	Yes	MM/SM	2x?	5, 7 & 10	10x FD

6.5 Activity Signal

This activity indication signal is completely controlled by the NVMe HDD device. This signal is meant to drive an LED for each device CRU/FRU in the storage system.

Req ID	Description												
ACTSIG-1	The system designer is required to redrive this signal with the appropriate driver specification required by the specific LED chosen for this function.												
ACTSIG-2	The two functions for this signal are defined as “Device Power” and “Device Activity”: ● Power: This function indicates the device has power and has no issues with its power regulation. Once the LED is “ON”, it shall either remain “ON” or blink at the Activity frequency unless the device determines power is no longer within its operating range. ● Activity: This function indicates if the device is being used. <table><tr><th>Function</th><th>LED State</th><th>Device State</th></tr><tr><td>Power ON</td><td>LED is “ON”</td><td>Device is powered, no activity occurring</td></tr><tr><td>Activity</td><td>4 Hz nominal “blink” rate</td><td>Device is powered, activity occurring</td></tr><tr><td>Power OFF</td><td>LED is “OFF”</td><td>Device is not powered</td></tr></table>	Function	LED State	Device State	Power ON	LED is “ON”	Device is powered, no activity occurring	Activity	4 Hz nominal “blink” rate	Device is powered, activity occurring	Power OFF	LED is “OFF”	Device is not powered
Function	LED State	Device State											
Power ON	LED is “ON”	Device is powered, no activity occurring											
Activity	4 Hz nominal “blink” rate	Device is powered, activity occurring											
Power OFF	LED is “OFF”	Device is not powered											

6.6 System Design

Req ID	Description
SYSTM-1	Systems designed to support NVMe HDDs are not required to support SAS or SATA HDDs. Backwards compatibility for systems designed to support NVMe HDDs to also support SAS or SATA HDDs is optional, and up to the system designer.
SYSTM-2	If a SAS or SATA HDD is mistakenly plugged in a drive slot in a system designed to support NVMe HDDs, the signal definition and system design must ensure that no damage to the system or the device occurs.

7 Power

7.1 Connector Power/Colling Limits

TBD

8 Security

8.1 Security Requirements from OCP Datacenter NVMe SSD Specification

Req ID	Description
SEC-1	All requirements in the OCP Datacenter NVMe SSD Specification v2.0 Section 11 are requirements for NVMe HDDs with the exceptions stated in this table
SEC-2	SPDM-3 is not a requirement for NVMe HDDs (FW attestation over SMBus/I2C)
SEC-3	SPDM-5 is an optional requirement for NVMe HDDs, not a mandatory requirement (FW attestation over PCIe DOE)