

OPEN
Compute Project

NVMe HDD Specification

Revision 0.7a
November 21, 2022

Contributing Companies/Contact
Microsoft/Jason Adrian
Seagate Technologies/Mohamad El-Batal
Western Digital/Dave Landsman

Revision History

Version	Date	Details
0.1	7/28/2021	First draft
0.2	8/27/2021	9/2/2021 draft review
0.3	5/18/2022	Initial Release 5/19/2022 Review Draft
0.4	5/20/2022	Initial Public Release Draft Candidate
0.5	6/2/2022	Update NVMe reference doc revisions and features table
0.6	07/21/2022	Updated connector spec, activity signal, decoupling
0.7	08/30/2022	Draft 6: Updated formatting, updated reference documents; updated NVMe section; updated Security section; numerous editorial changes.
	09/27/2022 (dl)	Draft 6 Edits: ● Added Compliance w/ OCP Tenets ● Added appendices A-C
	10/06/2022 (dl)	Cleanup remaining comments and edited Appendices A-B
0.7a	11/21/2022	● Updated table of Figures page 4 ● Removed Measurement Methodology section; was TBD but is not required in design spec ● Removed the NVMe PCIe Transport section; was TBD but is captured with other NVMe reqt's and other sections ● Re-entered fields in Appendix A that got lost somehow ● Removed Appendix B & C ● Removed "reference" version of NVMe requirements section; this duplicated sections in the spec and will be saved separately

Table of Contents

1 OCP 5
1.1 License 5
1.2 Acknowledgement 5

2 Compliance w/ OCP Tenets 6
2.1 Openness 6
2.2 Efficiency 6
2.3 Impact 7
2.4 Scale 7

3 Scope 7

4 References 8

5 Mechanical Form Factor 9
5.1 Mechanical Dimensions 9
5.2 Drive Side Label 9
5.3 Drive Top Label 9

6 Electrical Interface 10
6.1 Supported Connectors 10
6.1.1 Connector Signal Integrity Measured Data (Informative) 10
6.2 Connector AC Coupling 12
6.3 Single Port Connector (System / Device) 13
6.4 Dual Port Connector (System / Device) 14
6.5 PWRDIS 15
6.6 Activity Signal 15
6.7 System Design 15

7 Power 16
7.1 Connector Power Limits (Informative) 16
7.2 Device Power Limits 16
7.2.1 Max Continuous Current during Device Spin-up 16
7.2.2 Max Peak Current 16

8 PCIe 17

9 NVMe 18
9.1 NVMe Overview 18
9.2 NVMe Base Specification 18
9.3 NVMe NVM Command Set Specification 26
9.4 NVMe Zoned Namespace Command Set Specification 29
9.5 NVMe Management Interface Specification 30

10 Security 34

11 For future consideration 34

APPENDIX A - IC APPROVAL CHECKLIST **34**

Table of Figures

Figure 1: Label Content Requirements 8
Figure 2: Amphenol Measured Electrical Signal Integrity Graphs 11
Figure 3: PCIe Required AC-Coupling for NVMe Host & Device 12

1 OCP

1.1 License

Contributions to this Specification are made under the terms and conditions set forth in Open Web Foundation Modified Contributor License Agreement (“OWF CLA 1.0”) (“Contribution License”) by:

Microsoft
Seagate
Western Digital

Usage of this Specification is governed by the terms and conditions set forth in Open Web Foundation Modified Final Specification Agreement (“OWFa 1.0”) (“Specification License”). You can review the applicable OWFa1.0 Specification License(s) referenced above by the contributors to this Specification on the OCP website at http://www.opencompute.org/participate/legal-documents/.

For actual executed copies of either agreement, please contact OCP directly.

Notes: 1) The above license does not apply to the Appendix or Appendices. The information in the Appendix or Appendices is for reference only and non-normative in nature.

NOTWITHSTANDING THE FOREGOING LICENSES, THIS SPECIFICATION IS PROVIDED BY OCP "AS IS" AND OCP EXPRESSLY DISCLAIMS ANY WARRANTIES (EXPRESS, IMPLIED, OR OTHERWISE), INCLUDING IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, FITNESS FOR A PARTICULAR PURPOSE, OR TITLE, RELATED TO THE SPECIFICATION. NOTICE IS HEREBY GIVEN, THAT OTHER RIGHTS NOT GRANTED AS SET FORTH ABOVE, INCLUDING WITHOUT LIMITATION, RIGHTS OF THIRD PARTIES WHO DID NOT EXECUTE THE ABOVE LICENSES, MAY BE IMPLICATED BY THE IMPLEMENTATION OF OR COMPLIANCE WITH THIS SPECIFICATION. OCP IS NOT RESPONSIBLE FOR IDENTIFYING RIGHTS FOR WHICH A LICENSE MAY BE REQUIRED IN ORDER TO IMPLEMENT THIS SPECIFICATION. THE ENTIRE RISK AS TO IMPLEMENTING OR OTHERWISE USING THE SPECIFICATION IS ASSUMED BY YOU. IN NO EVENT WILL OCP BE LIABLE TO YOU FOR ANY MONETARY DAMAGES WITH RESPECT TO ANY CLAIMS RELATED TO, OR ARISING OUT OF YOUR USE OF THIS SPECIFICATION, INCLUDING BUT NOT LIMITED TO ANY LIABILITY FOR LOST PROFITS OR ANY CONSEQUENTIAL, INCIDENTAL, INDIRECT, SPECIAL OR PUNITIVE DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND EVEN IF OCP HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

1.2 Acknowledgement

The NVMe HDD specification was created under a collaboration from many OCP member companies and facilitated by the NVMe HDD Subgroup under the OCP Storage Workgroup. The NVMe HDD Subgroup would like to acknowledge the following member companies for their contributions to the NVMe HDD specification:

OCP Member Company
Amphenol Corporation
ARM
Broadcom Limited
Kioxia
Marvell Semiconductor, Inc.
Meta
Microchip Technology
Microsoft
Molex
Quanta
Seagate Technology
Super Micro
TE Connectivity Corporation
Toshiba
Western Digital
Wiwynn

2 Compliance w/ OCP Tenets

2.1 Openness

The NVMe HDD specification was developed in dialogue with the entire OCP Storage Project membership and the resulting specification represents an initial requirements set for this new category of device. The ecosystem is nascent, so we know the specification will evolve, but this spec represents the wisdom of the collective ecosystem on what are the requirements for building an NVMe HDD.

2.2 Efficiency

The efficiency created by this specification is that drive manufactures across the industry can build NVMe HDD devices to a consistent set of requirements, vetted by vendors and customers. Architecturally, NVMe HDD brings the NVMe/PCIe infrastructure to HDD's, enabling datacenter storage designers to take advantage of using the same storage infrastructure of HDD's that they do for SSDs. Additionally, NVMe/PCIe enables P2P operations to/between HDD, SSD, and compute accelerator layers using consistent protocols across the datacenter. Lastly, the adoption of NVMe enables scaling through NVMe over Fabrics

2.3 Impact

NVMe HDDs will have a huge impact on the storage ecosystem, bringing NVMe/PCIe storage devices with rotating media into the same storage protocol ecosystem as SSDs and compute accelerators. The scale of the HDD business speaks for itself in this regard.

2.4 Scale

There is no sales data for NVMe HDD but the expectation is that this segment will become a major part of the storage market

3 Scope

This specification defines the key attributes of a NVMe Hard Disk Drive.

4 References

From https://www.pcisig.com:
- PCI Express® Base Specification, Revision 4.0
- PCI Express® Base Specification, Revision 3.0
- 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://www.snia.org/:
- SFF-8482 Serial Attachment 2X Unshielded Connector, Revision 2.5
- SFF-8639 Multifunction 6X Unshielded Connector Revision 2.1
- SFF-8301 Specification for 3.5" Form Factor Drive Dimensions, Revision 1.9
- SFF-8323 Specification for 3.5" Form Factor Drive with Serial Attached Connector, Revision 1.6

From https://www.opencompute.org/:
- Datacenter NVMe® SSD Specification, Revision 2.0

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

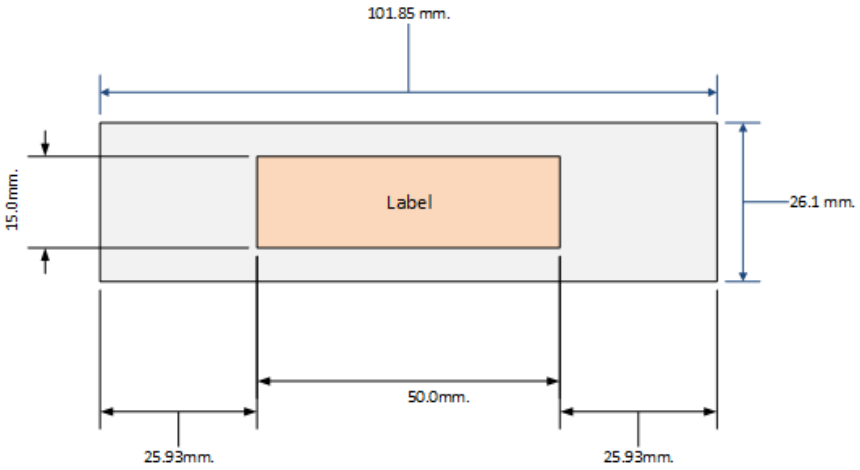
From https://serialata.org
- Serial ATA (SATA) Revision 3.5a

5 Mechanical Form Factor

5.1 Mechanical Dimensions

Req ID	Description
MECH-1	Physical form-factor mechanical dimensions must adhere to SFF-8301 Specification for 3.5" Form Factor Drive Dimensions (Rev 1.9)
MECH-2	Connector location must adhere to SFF-8323 Specification for 3.5" Form Factor Drive with Serial Attached Connector (Rev 1.6)

5.2 Drive Side Label

Req ID	Description
SDLBL-1	The top side of the NVMe HDD, the surface opposite the connector side of the drive, must have a defined label location to provide the drive part number and serial number. Dimensions and location are defined in Figure 1.  *Figure 1: Label Content Requirements*
SDLBL-2	Drive side label must contain the following; other details optional and vendor defined. 1) Drive Part number, both in text and 1D barcode 2) Drive serial number, both in text and 1D barcode Optional: 2D barcode

5.3 Drive Top Label

Req ID	Description
TPLBL-1	NVMe HDD drive top label must comply with the label requirements in the OCP Datacenter NVMe SSD Specification Version 2.0

6 Electrical Interface

6.1 Supported Connectors

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>CNNCT-1</td><td>The system connector PCIe pins must support PCIe Gen 3 signaling rates at a minimum, with recommended signaling rates meeting PCIe Gen 4.</td></tr>
<tr><td>CNNCT-2</td><td>NVMe HDDs (Host and Device) must support PCIe SRIS. There is no provision on the NVMe HDD to support REFCLK.</td></tr>
<tr><td>CNNCT-3</td><td>Supported system connector receptable / device connector plug combinations:
<table>
<tr><th>System Connector Receptacle</th><th>Device Connector Plug</th><th>Max PCIe link speed (8/16GT)</th></tr>
<tr><td>6Gb SATA</td><td>6Gb SATA</td><td></td></tr>
<tr><td>SFF8482 - 12Gb SAS</td><td>6Gb SATA</td><td>8Gb/s</td></tr>
<tr><td>SFF8482 - 24Gb SAS</td><td>6Gb SATA</td><td></td></tr>
<tr><td>SFF8482 - 12Gb SAS</td><td>12Gb SAS</td><td>8Gb/s</td></tr>
<tr><td>SFF8482 - 24Gb SAS</td><td>24Gb SAS</td><td>8Gb/s & 16Gb/s</td></tr>
<tr><td>SFF8482 - 24Gb SAS</td><td>12Gb SAS</td><td></td></tr>
<tr><td>SFF8639 - Gen4</td><td>6Gb SATA</td><td></td></tr>
<tr><td>SFF8639 - Gen4</td><td>12Gb SAS</td><td>8Gb/s</td></tr>
<tr><td>SFF8639 - Gen4</td><td>24Gb SAS</td><td>8Gb/s & 16Gb/s</td></tr>
</table>
</td></tr>
<tr><td>CNNCT-4</td><td>There is no new mechanical connector keying for NVMe HDDs. NVMe HDDS leverage existing connector keying for SAS and SATA connectors.
Note: The SAS system connector receptacle is typically capable of supporting a SAS or SATA device connector plug.</td></tr>
</table>

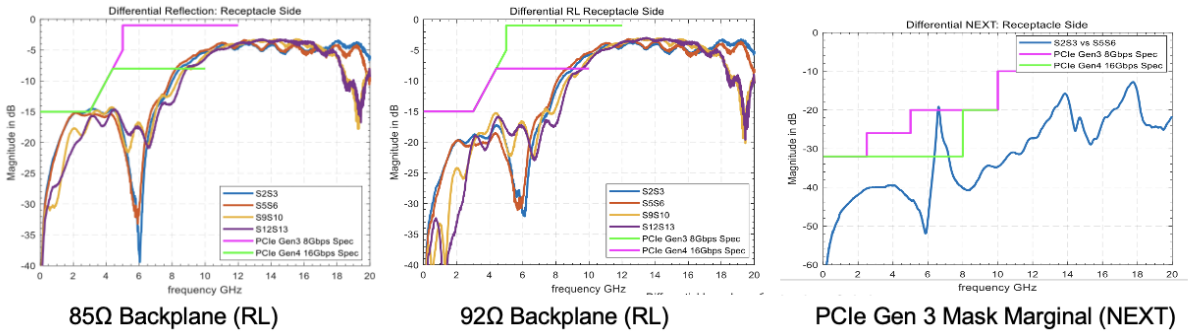
6.1.1 Connector Signal Integrity Measured Data (Informative)

The OCP NVMe HDD working group accepted the following Amphenol connector SI results:

- SAS3 Plug + SAS3 Receptacle for PCIe Gen 3
- SATA Plug + SAS3 Receptacle for PCIe Gen 3
- SAS4 Plug + SAS4 Receptacle for both PCIe Gen3 & Gen 4

Conn-Pair-Combination	PCIe Gen 3				PCIe Gen 4			
SI Test Type	**IL**	**RL**	**NEXT**	**FEXT**	**IL**	**RL**	**NEXT**	**FEXT**
SATA P/R-Pair	P	F	P	N/A	-	-	-	-
SAS3 P/R-Pair	**P**	**P/M***	**P**	**N/A**	P	F	P	P
SATA-P + SAS3-R Pair	**P**	**P**	**P/M****	**N/A**	P	F	F	N/A
SATA-P + SAS4-R Pair	**P**	**P**	**P**	**P**	F	F	F	N/A
SAS4 P/R-Pair	**P**	**P**	**P**	**P**	**P**	**P**	**P**	**P**
SAS3-P + SAS4-R Pair	**P**	**P**	**P**	**P**	F	F	P	P

IL = Insertion Loss, RL = Return Loss, NEXT = Near End Crosstalk, FEXT = Far End Crosstalk
P = Passed, M = Marginal, F = Failed
*P/M – Passed for frequencies under the 1.5x the base 4GHz frequency of PCIe-Gen3; however, Amphenol OCP representative considers it “Marginal” due to the fact that the Near-End-Crosstalk (NEXT) plot slightly exceeds the PCIe Gen3 around ~6.5MHz.
**P/M – Passed Return Loss (RL) with 92Ω system backplane PCB impedance, but Marginal with system 85Ω backplane PCB impedance.



85Ω Backplane (RL) | 92Ω Backplane (RL) | PCIe Gen 3 Mask Marginal (NEXT)

Figure 2: Amphenol Measured Electrical Signal Integrity Graphs

6.2 Connector AC Coupling

Req ID	Description
CNNAC-1	AC-coupling of Tx and Rx signals must be as shown in Figure 3 with capacitors values as in Figure 3 and Table 5:

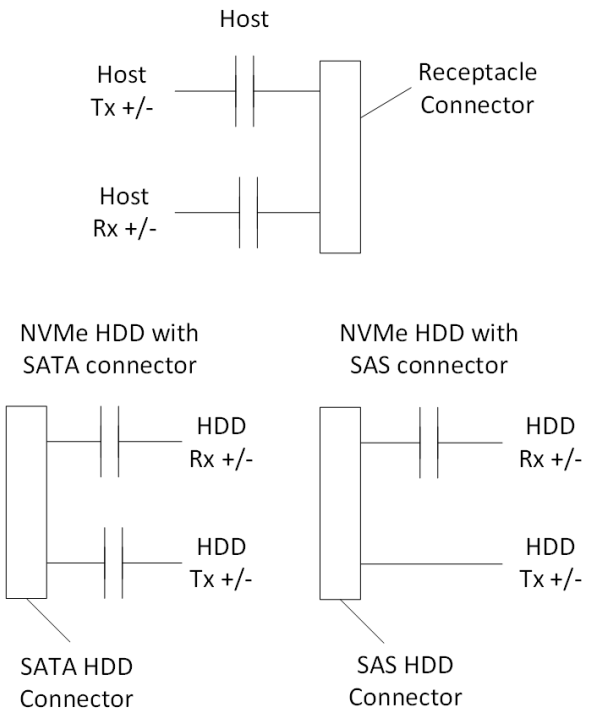


Figure 3: PCIe Required AC-Coupling for NVMe Host & Device

Location	Minimum	Maximum
Host	176 nF	265 nF
Single port NVMe HDD	NA	12 nF
Dual port NVMe HDD	176 nF	265 nF

6.3 Single Port Connector (System / Device)

The following Connector Pinout was taken directly from the SATA 3.5a specification to allow for using standard mass-produced SATA HDD SKUs and converted to Single-Port/Lane NVMe HDD SKUs with different firmware only.

Req ID	Description

SPPINN-1

System Side		Device Side	
p1	Reserved	p1	connect p2
p2	Reserved	p2	connect p1
p3	PWDIS - Power Disable	p3	PWDIS - Power Disable
p4	Ground	p4	Ground
p5	Ground	p5	Ground
p6	Ground	p6	Ground
p7	5v	p7	5v
p8	5v	p8	5v
p9	5v	p9	5v
p10	Ground	p10	Ground
p11	Activity – Input	p11	Activity – Driver
p12	Ground	p12	Ground
p13	12v	p13	12v
p14	12v	p14	12v
p15	12v	p15	12v
s1	Ground	s1	Ground
s2	TX+ (Host to Drive)	s2	RX+ (Drive from Host)
s3	TX- (Host to Drive)	s3	RX- (Drive from Host)
s4	Ground	s4	Ground
s5	RX- (Host from Drive)	s5	TX- (Drive to Host)
s6	RX+ (Host from Drive)	s6	TX+ (Drive to Host)
s7	Ground	s7	Ground

6.4 Dual Port Connector (System / Device)

The following Connector Pinout was taken directly from the SAS spec to allow for using standard mass-produced SAS HDD SKUs and converted to Dual-Port/Lane NVMe HDD SKUs with different firmware only.

Req ID	Description

DPPINN-1

System Side		Device Side	
p1	Reserved	p1	connect p2
p2	Reserved	p2	connect p1
p3	PWDIS - Power Disable	p3	PWDIS - Power Disable
p4	Ground	p4	Ground
p5	Ground	p5	Ground
p6	Ground	p6	Ground
p7	5v	p7	5v
p8	5v	p8	5v
p9	5v	p9	5v
p10	Ground	p10	Ground
p11	Activity – Input	p11	Activity – Driver
p12	Ground	p12	Ground
p13	12v	p13	12v
p14	12v	p14	12v
p15	12v	p15	12v
s1	Ground	s1	Ground
s2	P0-TX+ (Host to Drive)	s2	P0-RX+ (Drive from Host)
s3	P0-TX- (Host to Drive)	s3	P0-RX- (Drive from Host)
s4	Ground	s4	Ground
s5	P0-RX- (Host from Drive)	s5	P0-TX- (Drive to Host)
s6	P0-RX+ (Host from Drive)	s6	P0-TX+ (Drive to Host)
s7	Ground	s7	Ground
s8	Ground	s8	Ground
s9	P1-TX+ (Host to Drive)	s9	P1-RX+ (Drive from Host
s10	P1-TX- (Host to Drive)	s10	P1-RX- (Drive from Host)
s11	Ground	s11	Ground
s12	P1-RX- (Host from Drive)	s12	P1-TX- (Drive to Host)
s13	P1-RX+ (Host from Drive)	s13	P1-TX+ (Drive to Host)
s14	Ground	s14	Ground

6.5 PWRDIS

Req ID	Description
PWRDIS-1	The PWRDIS signal is asserted by the host to command the device to shut off power to all circuitry connected to the 12V and 5V power supplies. It must be supported by the device and is optional for the host.

6.6 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.

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

6.7 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 Limits (Informative)

	Voltage	Current per Pin	Number of Pins	Total Current
Max Continuous Current	+5V	1.5A	2	3.0A
	+12V	2.5A	2	5.0A
Max Peak Current	+5V	1.5A	2	3.0A
	+12V	2.5A	2	5.0A

7.2 Device Power Limits

7.2.1 Max Continuous Current during Device Spin-up

- Max for +12V → 2.35A
- Max for +5V → 1.4A

7.2.2 Max Peak Current

The NVMe HDD Max Peak currents must not exceed connector spec

8 PCIe

Req ID	Description
PCIE-1	The minimum number of ports is 1. The maximum number of ports is 2.
PCIE-2	The number of lanes per port: Single port: 1 or 2 lanes per port Dual port: 1 lane per port
PCIE-3	The minimum number of physical functions: 1
PCIE-4	The minimum number of virtual functions: 0
PCIE-5	The minimum PCIe base spec revision: 3.0
PCIE-6	The minimum link speed: TBD
PCIE-7	PCIe Link power management shall be supported
PCIE-8	PCIe Slot Power Control shall be supported

9 NVMe

9.1 NVMe Overview

Req ID	Description
NVME-1	NVMe HDDs shall be compatible with the following NVMe specifications as defined in this document: ● NVMe Base Specification ● NVMe NVM Command Set Specification ● NVMe PCIe Transport Specification ● NVMe Management Interface Specification
NVME-2	HDDs utilizing Shingled Magnetic Recording (SMR) shall be compatible with the following NVMe specifications as defined in this document: ● NVMe Zoned Namespace Command Set Specification
NVME-3	Support of the NVMe Key Value Command Set Specification is optional
NVME-4	Support of the NVMe Zoned Namespace Command Set Specification is: - Optional for CMR devices - Mandatory for SMR devices
NVME-5	HDDs utilizing multiple actuators shall represent the individual actuators as NVMe Namespaces

9.2 NVMe Base Specification

Req ID	Description
NBASE-1	NVMe HDDs shall support all Mandatory commands and capabilities in the NVMe Base Specification (version 2.0 or later) for an NVM subsystem with an I/O controller connected to a memory-based / PCIe-based transport.
NBASE-2	.The NVM subsystem shall have at least one namespace. Namespaces may be shared or private.
NBASE-3	The NVMe subsystem shall have a minimum of one controller of I/O controller type. Virtualization Enhancements and the Virtualization Management command support is optional.

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>NBASE-4</td><td>NVMe Base Spec admin command support:

<table>
<tr><th>Command</th><th>NVMe Spec Base 2.0 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>Delete I/O Submission Queue</td><td>M</td><td>M</td></tr>
<tr><td>Create I/O Submission Queue</td><td>M</td><td>M</td></tr>
<tr><td>Get Log Page</td><td>M</td><td>M</td></tr>
<tr><td>Delete I/O Completion Queue</td><td>M</td><td>M</td></tr>
<tr><td>Create I/O Completion Queue</td><td>M</td><td>M</td></tr>
<tr><td>Identity</td><td>M</td><td>M</td></tr>
<tr><td>Abort</td><td>M</td><td>M</td></tr>
<tr><td>Set Feature</td><td>M</td><td>M</td></tr>
<tr><td>Get Feature</td><td>M</td><td>M</td></tr>
<tr><td>Asynchronous Event Request</td><td>M</td><td>M</td></tr>
<tr><td>Capacity Management</td><td>O</td><td>O</td></tr>
<tr><td>Namespace Management</td><td>O</td><td>O</td></tr>
<tr><td>Firmware Commit</td><td>O</td><td>M</td></tr>
<tr><td>Firmware Image Download</td><td>O</td><td>M</td></tr>
<tr><td>Device Self-test</td><td>O</td><td>M</td></tr>
<tr><td>Namespace Attachment</td><td>O</td><td>O</td></tr>
<tr><td>Keep Alive</td><td>O</td><td>O</td></tr>
<tr><td>Directive Send</td><td>O</td><td>O</td></tr>
<tr><td>Directive Receive</td><td>O</td><td>O</td></tr>
<tr><td>Virtualization Management</td><td>O</td><td>O</td></tr>
<tr><td>NVMe-MI Send</td><td>O</td><td>M</td></tr>
<tr><td>NVMe-MI Receive</td><td>O</td><td>M</td></tr>
<tr><td>Doorbell Buffer Config</td><td>O</td><td>O</td></tr>
<tr><td>Lockdown</td><td>O</td><td>M</td></tr>
<tr><td>Format NVM</td><td>O</td><td>O</td></tr>
<tr><td>Security Send</td><td>O</td><td>M</td></tr>
<tr><td>Security Receive</td><td>O</td><td>M</td></tr>
<tr><td>Sanitize</td><td>O</td><td>M</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
</table>

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>NBASE-5</td><td>NVMe Base Spec common I/O command support:

<table>
<tr><th>Command</th><th>NVMe Spec Base 2.0 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>Flush</td><td>M</td><td>M</td></tr>
<tr><td>Dataset Management (primarily in NVM Command Set spec)</td><td>O</td><td>O</td></tr>
<tr><td>Reservation Register</td><td>O [1]</td><td>O [1]</td></tr>
<tr><td>Reservation Report</td><td>O [1]</td><td>O [1]</td></tr>
<tr><td>Reservation Acquire</td><td>O [1]</td><td>O [1]</td></tr>
<tr><td>Reservation Release</td><td>O [1]</td><td>O [1]</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited

NOTES
[1] - Mandatory if reservations are supported</td></tr>
<tr><td>NBASE-6</td><td>SGL support is optional.</td></tr>
<tr><td>NBASE-7</td><td>Queues:
1) The minimum number of I/O submission/completion queue pairs per controller is 4.
2) The minimum I/O submission/completion queue depth is TBD.
3) The minimum Admin submission/completion queue depth is TBD.</td></tr>
</table>

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>NBASE-8</td><td>NVMe Base Spec log page support:

<table>
<tr><th>Log Page</th><th>NVMe Spec Base 2.0 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>Supported Log Pages</td><td>M</td><td>M</td></tr>
<tr><td>Error Information</td><td>M</td><td>M</td></tr>
<tr><td>SMART / Health Information (Controller scope)</td><td>M</td><td>M</td></tr>
<tr><td>SMART / Health Information (Namespace scope)</td><td>O</td><td>O</td></tr>
<tr><td>Firmware Slot Information</td><td>O</td><td>M</td></tr>
<tr><td>Changed Namespace List</td><td>O</td><td>M</td></tr>
<tr><td>Commands Supported and Effects</td><td>M</td><td>M</td></tr>
<tr><td>Device Self-test</td><td>O</td><td>M</td></tr>
<tr><td>Telemetry Host-Initiated</td><td>O</td><td>M</td></tr>
<tr><td>Telemetry Controller-Initiated</td><td>O</td><td>M</td></tr>
<tr><td>Endurance Group Information</td><td>O</td><td>M</td></tr>
<tr><td>Predictable Latency Per NVM Set</td><td>O</td><td>O</td></tr>
<tr><td>Predictable Latency Event Aggregate</td><td>O</td><td>O</td></tr>
<tr><td>Asymmetric Namespace Access</td><td>O</td><td>O</td></tr>
<tr><td>Persistent Event</td><td>O</td><td>O</td></tr>
<tr><td>Endurance Group Event Aggregate</td><td>O</td><td>M</td></tr>
<tr><td>Media Unit Status</td><td>O [1]</td><td>O [1]</td></tr>
<tr><td>Supported Capacity Configuration List</td><td>O [1]</td><td>O [1]</td></tr>
<tr><td>Feature Identifiers Supported and Effects</td><td>M</td><td>M</td></tr>
<tr><td>NVMe-MI Commands Supported and Effects</td><td>M</td><td>M</td></tr>
<tr><td>Command and Feature Lockdown</td><td>O</td><td>M</td></tr>
<tr><td>Reservation Notification</td><td>O</td><td>O</td></tr>
<tr><td>Sanitize Status</td><td>O</td><td>M</td></tr>
<tr><td>Boot Partition</td><td>O</td><td>O</td></tr>
<tr><td>Physical Interface Receiver Eye Opening Measurement</td><td>O</td><td>See TODO-1</td></tr>
<tr><td>Rotational Media Information</td><td>O</td><td>M</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited
NOTES
[1] - Mandatory for controllers that support Fixed Capacity Management</td></tr>
</table>

NBASE-9

NVMe Base Spec feature support:

Feature	NVMe Spec Base 2.0 Support (for reference only)	OCP NVMe HDD Support
Arbitration	M	M
Power Management	M	M
Temperature Threshold	M	M
Volatile Write Cache	O	M
Number of Queues	M	M
Interrupt Coalescing	M	M
Interrupt Vector Configuration	M	M
Asynchronous Event Configuration	O	M
Autonomous Power State Transition	O	O
Host Memory Buffer	O	O
Timestamp	O	M
Keep Alive Timer	O	O
Host Controlled Thermal Management	O	O
Non-Operational Power State Config	O	O
Read Recovery Level Config	O	O
Predictable Latency Mode Config	O	O
Predictable Latency Mode Window	O	O
Host Behavior Support	O	O
Sanitize Config	O	M
Endurance Group Event Configuration	O	M
I/O Command Set Profile	O	O [2]
Software Progress Marker	O	O
Host Identifier	O [1]	O [1]
Reservation Notification Mask	O [1]	O [1]
Reservation Persistence	O [1]	O [1]
Namespace Write Protection Config	O	O
Enhanced Controller Metadata	See NVMe-MI	O
Controller Metadata	See NVMe-MI	O
Namespace Metadata	See NVMe-MI	O
Spinup Control (Rotational Media)	O	M

O/M/P/NR definition: O = Optional, M = Mandatory, P = Prohibited, NR = Not Recommended

NOTES

[1] – Mandatory if reservations are supported

[2] – Mandatory for NVMe HDDs supporting not just the NVM Command Set, i.e. NVMe HDDs supporting ZNS or Key Value, Optional for conventional NVMe HDDs

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>NBASE-10</td><td>The ‘Weighted Round Robin with Urgent Priority Class Arbitration’ method for command arbitration shall be supported</td></tr>
<tr><td>NBASE-11</td><td>NVMe Base Spec extended capabilities support:
<table>
<tr><th>Extended Capability</th><th>NVMe Spec Base 2.0 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>Asymmetric Namespace Access Reporting</td><td>O</td><td>O</td></tr>
<tr><td>Boot Partitions</td><td>O</td><td>O</td></tr>
<tr><td>Capacity Management</td><td>O</td><td>O</td></tr>
<tr><td>Command and Feature Lockdown</td><td>O</td><td>M</td></tr>
<tr><td>Controller Memory Buffer</td><td>O</td><td>M</td></tr>
<tr><td>Device Self-test Operations</td><td>O</td><td>M</td></tr>
<tr><td>Directives</td><td>O</td><td>O</td></tr>
<tr><td>Doorbell Stride for Software Emulation</td><td>O</td><td>O</td></tr>
<tr><td>Host Memory Buffer</td><td>O</td><td>O</td></tr>
<tr><td>Namespace Management</td><td>O</td><td>O</td></tr>
<tr><td>Namespace Write Protection</td><td>O</td><td>O</td></tr>
<tr><td>Persistent Memory Region</td><td>O</td><td>O</td></tr>
<tr><td>Power Management</td><td>see NBASE-9</td><td>see NBASE-9 and NBASE-15</td></tr>
<tr><td>Predictable Latency Mode</td><td>O</td><td>O</td></tr>
<tr><td>Read Recovery Level</td><td>O</td><td>O</td></tr>
<tr><td>Replay Protected Memory Block</td><td>O</td><td>O</td></tr>
<tr><td>Reservations</td><td>O</td><td>O</td></tr>
<tr><td>Rotational Media</td><td>O</td><td>M</td></tr>
<tr><td>Sanitize Operations</td><td>O</td><td>M</td></tr>
<tr><td>Submission Queue (SQ) Associations</td><td>O</td><td>O</td></tr>
<tr><td>Telemetry</td><td>O</td><td>M</td></tr>
<tr><td>Virtualization Enhancements</td><td>O</td><td>O</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
<tr><td>NBASE-12</td><td>The size of the CMB shall be a minimum of 16MiB</td></tr>
<tr><td>NBASE-13</td><td>The CMB shall support Write Data (WDS), Read Data (RDS), PRP SGL Lists (LISTS), Completion Queues (CQS) and Submission Queues (SQS).</td></tr>
<tr><td>NBASE-14</td><td>Supported sanitize operation types shall include Overwrite and Crypto Erase.</td></tr>
</table>

Req ID	Description
NBASE-15	NVMe Power management details 1) Identify controller (CNS=01h) shall report at least these power states, modeled after the SATA Extended Power Conditions states, in descending order of power use. 0: corresponding to the spinup process / operational (e.g., more power used than normal for a short duration) 1: corresponding to the normal operational state 2: corresponding to a reduced power state (IdleA); operational (e.g., spinning, heads tracking) 3: corresponding to a further reduced power state (IdleB); operational (e.g., spinning, heads unloaded) 4: corresponding to the lowest power state (Standby); non-operational (e.g., not spinning) 2) The Set/Get Features (Feature ID=Power Management) commands are used by the host to change NVMe power states. 3) Any media access command shall transition to the normal operational state from any other power state.

9.3 NVMe NVM Command Set Specification

Req ID	Description
NVMCS-1	NVMe HDDs shall support all Mandatory commands and capabilities in the NVMe NVM Command Set Specification (version 1.0 or later) for an NVM subsystem with an I/O controller.
NVMCS-2	NVMe NVM Command Set Spec admin command support: Command / NVMe Spec NVM Cmd Set 1.0 Support (for reference only) / OCP NVMe HDD Support Get LBA Status / O / O O/M/P definition: O = Optional, M = Mandatory, P = Prohibited
NVMCS-3	For all supported admin commands from the NVMe Base specification, any NVM Command Set specific behavior shall be supported.
NVMCS-4	LB size of 4KiB shall be supported. All other LB sizes are optional. Support of Metadata, including PI, is optional.

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>NVMCS-5</td><td>NVMe NVM Command Set Spec NVM command set support:

<table>
<tr><th>Command</th><th>NVMe Spec NVM Cmd Set 1.0 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>Write</td><td>M</td><td>M</td></tr>
<tr><td>Read</td><td>M</td><td>M</td></tr>
<tr><td>Write Uncorrectable</td><td>O</td><td>O</td></tr>
<tr><td>Compare</td><td>O</td><td>O</td></tr>
<tr><td>Copy</td><td>O</td><td>O</td></tr>
<tr><td>Write Zeroes</td><td>O</td><td>O</td></tr>
<tr><td>Verify</td><td>O</td><td>O</td></tr>
<tr><td>Dataset Management (also referenced in Base Spec)</td><td>O</td><td>O</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
<tr><td>NVMCS-6</td><td>Support for Fused Operation of Compare and Write commands is optional.</td></tr>
<tr><td>NVMCS-7</td><td>NVMe NVM Command Set Spec log page support:

<table>
<tr><th>Log Page</th><th>NVMe Spec NVM Cmd Set 1.0 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>LBA Status Information</td><td>O</td><td>O</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
<tr><td>NVMCS-8</td><td>For all supported log pages from the NVMe Base specification, any NVM Command Set specific behavior shall be supported.</td></tr>
<tr><td>NVMCS-9</td><td>NVMe NVM Command Set Spec feature support:

<table>
<tr><th>Feature</th><th>NVMe Spec NVM Cmd Set 1.0 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>LBA Range Information</td><td>O</td><td>O</td></tr>
<tr><td>Error Recovery</td><td>M</td><td>M</td></tr>
<tr><td>Write Atomicity Normal</td><td>M</td><td>M</td></tr>
<tr><td>LBA Status Information Attributes</td><td>O</td><td>O</td></tr>
</table>
O/M/P/NR definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
<tr><td>NVMCS-10</td><td>For all supported features from the NVMe Base specification, any NVM Command Set specific behavior shall be supported.</td></tr>
</table>

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>NVMCS-11</td><td>NVMe NVM Command Set Spec extended capabilities support:

<table>
<tr><th>Extended Capability</th><th>NVMe Spec NVM Cmd Set 1.0 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>End-to-end Data Protection</td><td>O</td><td>O</td></tr>
<tr><td>Get LBA Status</td><td>O</td><td>See NVMCS-2</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
<tr><td>NVMCS-12</td><td>For all supported extended capabilities from the NVMe Base specification, any NVM Command Set specific behavior shall be supported.</td></tr>
<tr><td>NVMCS-13</td><td>Supported LBA formats

<table>
<tr><th>LBA size</th><th>PI Size</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>512</td><td>0</td><td>O</td></tr>
<tr><td>512</td><td>8</td><td>O</td></tr>
<tr><td>4096</td><td>0</td><td>M</td></tr>
<tr><td>4096</td><td>8</td><td>O</td></tr>
<tr><td>4096</td><td>16</td><td>O</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
</table>

9.4 NVMe Zoned Namespace Command Set Specification

The ZNSCS requirements are only applicable if SMR recording techniques are supported.

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>ZNSCS-1</td><td>Support the NVM Express Zoned Namespace Command Set Specification (version 1.1 or later)</td></tr>
<tr><td>ZNSCS-2</td><td>The Zone Append command is optional</td></tr>
<tr><td>ZNSCS-3</td><td>Zone Management Receive command subcommands

<table>
<tr><th>Subcommand</th><th>NVMe Spec ZNS Cmd Set 1.1 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>Report Zones</td><td>M</td><td>M</td></tr>
<tr><td>Extended Report Zones</td><td>See ZNSCS-8</td><td>O</td></tr>
</table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
</table>

<table>
<tr><td>ZNSCS-4</td><td>Zone Management Send command subcommands

<table><tr><th>Subcommand</th><th>NVMe Spec ZNS Cmd Set 1.1 Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr><tr><td>Close Zone</td><td>M</td><td>M</td></tr><tr><td>Finish Zone</td><td>M</td><td>M</td></tr><tr><td>Open Zone</td><td>M</td><td>M</td></tr><tr><td>Reset Zone</td><td>M</td><td>M</td></tr><tr><td>Offline Zone</td><td>O</td><td>O</td></tr><tr><td>Set Zone Descriptor Extension</td><td>See ZNSCS-8</td><td>O</td></tr><tr><td>Flush Explicit ZRWA Range</td><td>See ZNSCS-9</td><td>O</td></tr></table>
O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
<tr><td>ZNSCS-5</td><td>Reset Zone Recommended support is optional.</td></tr>
<tr><td>ZNSCS-6</td><td>Finish Zone Recommended support is optional.</td></tr>
<tr><td>ZNSCS-7</td><td>Zone Active Excursions support is optional.</td></tr>
<tr><td>ZNSCS-8</td><td>Zone Descriptor Extension support is optional.
This includes
- Zone Management Receive (Extended Report Zones)
- Zone Management Send (Set Zone Descriptor)</td></tr>
<tr><td>ZNSCS-9</td><td>Zone Random Write Area (ZRWA) support is optional.
This includes:
- Zone Management Send (Flush Explicit ZRWA Range)</td></tr>
<tr><td>ZNSCS-10</td><td>TCG security interactions are as specified in TCG SIIS</td></tr>
<tr><td>ZNSCS-11</td><td>Read-Only zone condition support is optional.</td></tr>
</table>

9.5 NVMe Management Interface Specification

Req ID	Description
NMI-1	Support the NVM Express Zoned Namespace Command Set Specification (version 1.1 or later)
NMI-2	In-band tunneling of NVMe-MI Command Set commands via the NVMe admin commands NVMe-MI Send and NVMe-MI Receive shall be supported.
NMI-3	Processing NVMe-MI Command Set commands via an out-of-band PCIe VDM interface shall be supported. This stack includes support for the relevant MCTP specs as defined by DMTF, including the MCTP PCIe VDM Transport Binding spec.

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>NMI-4</td><td>Management Interface Command Set

<table>
<tr><th>Subcommand</th><th>NVMe NVMe-MI 1.1 Spec Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>Configuration Get</td><td>M</td><td>M</td></tr>
<tr><td>Configuration Set</td><td>M</td><td>M</td></tr>
<tr><td>Controller Health Status Poll</td><td>M</td><td>M</td></tr>
<tr><td>Management Endpoint Buffer Read</td><td>O</td><td>O</td></tr>
<tr><td>Management Endpoint Buffer Write</td><td>O</td><td>O</td></tr>
<tr><td>NVM Subsystem Health Status Poll</td><td>M</td><td>M</td></tr>
<tr><td>Read NVMe-MI Data Structure</td><td>M</td><td>M</td></tr>
<tr><td>Reset</td><td>O</td><td>O</td></tr>
<tr><td>SES Receive</td><td>O</td><td>O</td></tr>
<tr><td>SES Send</td><td>O</td><td>O</td></tr>
<tr><td>Shutdown</td><td>O</td><td>O</td></tr>
<tr><td>VPD Read</td><td>M</td><td>M</td></tr>
<tr><td>VPD Write</td><td>O</td><td>P</td></tr>
</table>

O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
<tr><td>NMI-5</td><td>NVM Express Admin Command Set (out of band)
- See the NVMe-MI specification for details

<table>
<tr><th>Subcommand</th><th>NVMe NVMe-MI 1.1 Spec Support (for reference only)</th><th>OCP NVMe HDD Support</th></tr>
<tr><td>Get Log Page</td><td>M</td><td>M</td></tr>
<tr><td>Identify</td><td>M</td><td>M</td></tr>
<tr><td>Set Features</td><td>O</td><td>M</td></tr>
<tr><td>Get Features</td><td>M</td><td>M</td></tr>
<tr><td>Virtualization Management</td><td>O</td><td>O</td></tr>
<tr><td>Lockdown</td><td>O</td><td>O</td></tr>
<tr><td>Format NVM</td><td>O</td><td>O</td></tr>
<tr><td>Security Send</td><td>O</td><td>O</td></tr>
<tr><td>Security Receive</td><td>O</td><td>O</td></tr>
<tr><td>Sanitize</td><td>O</td><td>O</td></tr>
<tr><td>Get LBA Status</td><td>O</td><td>O</td></tr>
<tr><td colspan="3">Note: all other commands are prohibited for NVMe HDD</td></tr>
</table>

O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
</table>

<table>
<tr><th>Req ID</th><th>Description</th></tr>
<tr><td>NMI-6</td><td>The PCIe Command Set should not be supported.

<table>
<tr><th>Subcommand</th><th>NVMe NVMe-MI 1.1 Spec Support (for reference only)</th><th>NVMe HDD Support</th></tr>
<tr><td>PCIe Configuration Read</td><td>O</td><td>O</td></tr>
<tr><td>PCIe Configuration Write</td><td>O</td><td>O</td></tr>
<tr><td>PCIe Memory Read</td><td>O</td><td>O</td></tr>
<tr><td>PCIe Memory Write</td><td>O</td><td>O</td></tr>
<tr><td>PCIe I/O Read</td><td>O</td><td>O</td></tr>
<tr><td>PCIe I/O Write</td><td>O</td><td>O</td></tr>
</table>

O/M/P definition: O = Optional, M = Mandatory, P = Prohibited</td></tr>
</table>

10 Security

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 (FW attestation over SMBus/I2C) is optional, not mandatory
SEC-3	SPDM-5 (FW attestation PCIe DOE) is optional, not mandatory

11 For future consideration

Req ID	Description
TODO-1	Physical Interface Receiver Eye Opening Measurement log page
TODO-2	Depopulation/Repopulation feature
TODO-3	Command Duration Limit feature

Appendix A - IC Approval Checklist

Complete all the checklist items in the table with links to the section where it is described in this spec or an external document.

Item	Status or Details	Link to detailed explanation
Is this contribution entered into the OCP Contribution Portal?	Yes	
Was it approved in the OCP Contribution Portal?	Approved by members of the Storage Project; awaiting final IC approval	
Is there a Supplier(s) that is building a product based on this Spec?	N/A; this is a “design spec”	N/A for a design spec