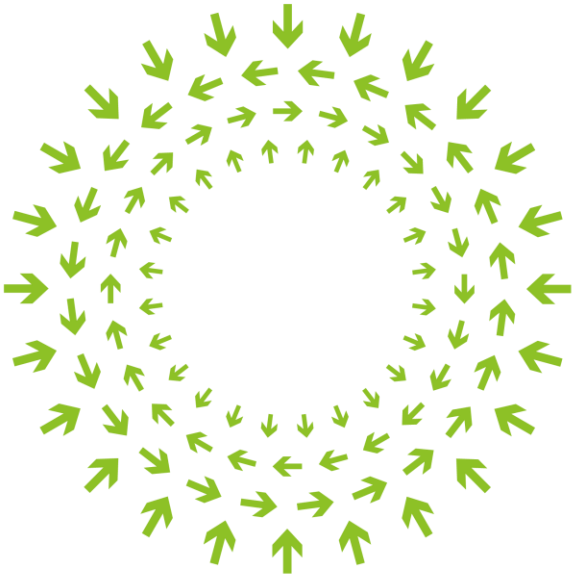


OPEN

Compute Project

Datacenter NVMe® SSD Specification

Version 2.0 (07302021)

Author: Ross Stenfort, Ta-Yu Wu, Facebook
Author: Lee Prewitt, Microsoft
Author: Paul Kaler, David Derosa, HPE
Author: William Lynn, Austin Bolen, Dell EMC

Table of Contents

1 OCP

1.1 License OWF Option

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

Facebook
Microsoft
HPE
Dell EMC

Usage of this Specification is governed by the terms and conditions set forth in the Open Web Foundation Final Specification Agreement ("OWFa 1.0").
Note: The following clarifications, which distinguish technology licensed in the Contribution License and/or Specification License from those technologies merely referenced (but not licensed), were accepted by the Incubation Committee of the OCP:

NONE

1.2 System Firmware

All products seeking OCP Accepted™ Product Recognition must complete the Open System Firmware (OSF) Tab in the 2021 Supplier Requirements Checklist. The completed checklist shall be uploaded and available
at: https://github.com/opencomputeproject/OpenSystemFirmware/[vendor_name]/[product_name]/

Note to author: replace [vendor_name] and [product_name] with actual company name and product identifier.

1.3 Hardware Management

1.3.1 Compliance

All products seeking OCP Inspired™ or OCP Accepted ™ Product Recognition shall comply with the OCP Hardware Management Baseline Profile V1.0 and provide such evidence by completing the Hardware Management Tab in the 2021 Supplier Requirements Checklist.

1.3.2 BMC Source Availability (if applicable)

All Products seeking OCP Accepted™ Product Recognition shall have source code and binary blobs submitted for BMC, if applicable. The BMC management source code shall be uploaded at: https://github.com/opencomputeproject/Hardware Management/[vendor_name]/[product_name]/.

1.4 Security

All products seeking OCP Inspired™ or OCP Accepted™ Product Recognition shall have a completed Security Profile in the 2021 Supplier Requirements Checklist. Whether the answer is a yes or no, the profile must be completed. For Additional Security Badges (Bronze/Silver/Gold), please fill out the Security Profile in accordance with the requirements for that level. Security Badges will be reassessed on an annual basis as requirements are subject to change.

2 Overview

This document is to define the requirements for a Datacenter NVMe SSD (DSSD) for use in datacenters.

3 Scope

This document covers PCIe-attached SSDs using NVM Express.

4 NVM Express Requirements

4.1 Overview

Requirement ID	Description
NVMe-1	The device shall comply with all required features of the NVMe 1.4b Specification. Optional features shall be implemented per the requirements of this specifications.
NVMe-2	Any optional features supported by the device not described in this document shall be clearly documented and disclosed to the customer.
NVMe-3	Any vendor unique features supported by the device not described in this document shall be clearly documented and disclosed to the customer.

4.2 NVMe Reset Supported

Requirement ID	Description
NVMeR-1	NVMe Subsystem Reset (NSSR) shall be supported.
NVMeR-2	NVMe Controller Reset (CC.EN cleared to 0b) shall be supported.

4.3 NVMe Controller Configuration and Behavior

Requirement ID	Description
NVMe-CFG-1	The default arbitration shall be Round-Robin. Weighted Round Robin with Urgent Class Priority shall be supported.
NVMe-CFG-2	The device shall support a Maximum Data Transfer Size (MDTS) value of at least 256KB.
NVMe-CFG-3	The device firmware shall support reporting of CSTS.CFS as indicated in the NVMe Specification version 1.4b.
NVMe-CFG-4	Obsolete. Replaced by LABL-11 (Model Number shall match).
NVMe-CFG-5	The minimum supported queue depth shall be 1024 per submission queue.
NVMe-CFG-6	The minimum number of I/O Queue Pairs shall be 64.

Requirement ID	Description
NVMe-CFG-7	Device shall support EUI64 (Extended Unique Identifier) to differentiate namespaces.
NVMe-CFG-8	Device shall support an NGUID per Namespace.

4.4 NVMe Admin Command Set

The device shall support the following mandatory and optional NVMe admin commands:

Requirement ID	Description
NVMe-AD-1	The device shall support all mandatory NVMe admin commands.
NVMe-AD-2	Identify – In addition to supporting all the mandatory CNS values and the associated mandatory fields within the CNS, the following optional fields in the CNS shall be supported: • Format progress indicator (FPI). • I/O Performance and Endurance Hints: ○ NSFEAT bit 4 = 1b.
NVMe-AD-3	Namespace Management command shall be supported.
NVMe-AD-4	Namespace Attachment command shall be supported.
NVMe-AD-5	Format NVM command shall be supported. Secure Erase Settings (SES) values 000b, 001b and 010b shall be supported.
NVMe-AD-6	The device shall support NVMe-MI Send and Receive.
NVMe-AD-7	The device shall support the Sanitize command and meet NIST SP800-88r1 Purge requirements. Block Erase (010b) and Crypto Erase (100b) sanitize operations shall be supported. If Overwrite (011b) is supported it shall meet NIST SP800-88r1 Purge requirements.
NVMe-AD-8	The device shall enable reads to sanitized LBAs to meet validation of sanitized areas per NIST SP800-88r1 (i.e., return data from media). Specifically, LBAs shall return all 0s or all 1s after a Block Erase (010b) and LBAs shall return the data generated from decryption using the new MEK after a Crypto Erase (100b) (garbage data).
NVMe-AD-9	If a read operation occurs to a Sanitized LBA prior to that LBA being written, the read operation shall not report a media error because the LBA has not been written.
NVMe-AD-10	The device shall support Identify command UUID List functionality (CNS value 17h).
NVMe-AD-11	If the Namespace Identifier (NSID) is not used for a specific NVMe admin command, and the host specifies a non-zero NSID, then the controller shall abort the command with status Invalid Field in Command.

4.4.1 Namespace Management/Attachment Commands

The namespace management command along with the attach/detach commands is used to increase device over-provisioning beyond the default minimum over-provisioning.

Requirement ID	Description
NSM-1	The namespace management commands shall be supported on all namespaces.
NSM-2	When creating a namespace, the default "Formatted LBA Size" parameter (FLBAS = 0) in the Identify Namespace Data Structure (Byte 26) shall correspond to the default sector size set at the factory.
NSM-3	When formatting the device with the Format command, the default "LBA Format" parameter (LBAF = 0) in Command Dword 10 bits 3:0 shall correspond to the default sector size set at the factory.
NSM-4	The device shall support a minimum of 16 Namespaces (see Section 12 Device Profiles).
NSM-5	For some models (see Section 12 Device Profiles), the minimum number of namespaces shall be at least 128 Namespaces at 7TB based on 16 Namespaces per TB of usable capacity. The number of Namespaces shall be based on the device usable capacity as follows:

Device Usable Capacity	Minimum Number of Namespaces
<= 1TB	16
> 1TB but <= 2TB	32
> 2TB but <= 3TB	48
> 3TB but <= 4TB	64
…	…
> 7TB	128

NSM-6	The device shall report at least one Namespace Granularity Descriptor in the Namespace Granularity List.
NSM-7	The device shall support the TNVMCAP and UNVMCAP fields in the Identify Controller Data Structure.

4.4.2 Namespace Utilization (NUSE)

Requirement ID	Description
NUSE-1	The NUSE shall be equal to the number of logical blocks currently allocated in the namespace. NUSE cannot be hardcoded to be equal to NCAP. Here is an example for a 200GB device:

Requirement ID	Description
	1. After a Format NVM command User Data Erase (SES = 001b), NUSE would be zero. And the usage data would reflect that: 0.00GB. 2. After writing 1 GB worth of data, the usage data is expected to show the following: 1.00GB. 3. After filling the device, the usage data is expected to show the following: 200.00GB. 4. If the host issues a 10GB de-allocate command and the device completes de-allocation of the data, the usage data would show the following: 190.00GB.

4.4.3 UUID for DSSD Specific Information

A UUID has been defined for use in commands to ensure that the vendor specific Log Identifiers and Feature Identifiers used in this specification access the functionality defined in this specification (i.e., do not access other vendor specific functionality that may use the same vendor specific identifiers).

Requirement ID	Description
UUID-1	The UUID List (NVMe-AD-10) shall contain a UUID List Entry that contains the UUID value C194D55BE0944794A21D29998F56BE6F.The Identifier Association field in that UUID List Entry shall be cleared to 00b.
UUID-2	The Get Features and Set Features commands shall support UUID Index functionality.
UUID-3	A Get Features command or a Set Features command with: • the UUID Index of the UUID (UUID-1) in the UUID List (NVMe-AD-10) or a zero UUID Index; and • a vendor-specific Feature Identifier that is used in this specification (see Section 4.12 Set/Get Feature Requirements) shall access the vendor specific Feature defined in this specification.
UUID-4	The Get Log Page command shall support UUID Index functionality.
UUID-5	A Get Log Page command with: • the UUID Index of the UUID (UUID-1) in the UUID List (NVMe-AD-10) or a zero UUID Index; and • a vendor-specific Log Page Identifier that is used in this specification (see Section 4.8 Log Page Requirements) shall access the vendor specific Log Page defined in this specification.

4.5 NVMe I/O Command Set

Requirement ID	Description
NVMe-IO-1	The device shall support all mandatory NVMe I/O commands.

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td>NVMe-IO-2</td><td>The device shall support the Dataset Management command. The device shall support the Attribute – Deallocate (AD) bit.</td></tr>
<tr><td>NVMe-IO-3</td><td>Since the device is power fail safe (e.g., has Power Loss Protection (PLP)) the performance shall not be degraded by any of the following:
FUA – i.e., forced unit access shall not incur a performance penalty.
Flush Cache – i.e., flush cache shall have no effect as the PLP makes any cache non-volatile.
Volatile Write Cache (Feature Identifier 06h) Set Feature to disable write-cache. This command shall be failed as described in the NVMe Standard 1.4b as there is no volatile write cache.</td></tr>
<tr><td>NVMe-IO-4</td><td>The device shall support the Write Zeroes command. The following bits of the Write Zeroes command shall be supported:
De-allocate (DEAC) bit.
Force Unit Access (FUA) bit.</td></tr>
<tr><td>NVMe-IO-5</td><td>The Write Zeroes command shall have the following behavior:
<table>
<tr><th>DEAC</th><th>FUA</th><th>Behavior</th></tr>
<tr><td>0b</td><td>0b</td><td>The device shall follow the NVMe 1.4b Specification.</td></tr>
<tr><td>0b</td><td>1b</td><td>The device shall follow the NVMe 1.4b Specification.</td></tr>
<tr><td>1b</td><td>1b</td><td>The device shall follow the NVMe 1.4b Specification.</td></tr>
<tr><td>1b</td><td>0b</td><td>See NVMe-IO-6 (Write Zeroes DEAC bit).</td></tr>
</table></td></tr>
<tr><td>NVMe-IO-6</td><td>If the Write Zeroes DEAC bit is set to 1b and the FUA bit is cleared to 0b, the device shall un-map the specified blocks and shall return a zero value for any subsequent read to the specified blocks regardless of the behavior of the Dataset Management command.</td></tr>
<tr><td>NVMe-IO-7</td><td>With the DEAC bit set to 1b and the FUA bit cleared to 0b one or more Write Zeros command(s) shall be able to completely update the FTL map of the entire device in less than one minute.</td></tr>
<tr><td>NVMe-IO-8</td><td>The device shall support the Compare command.</td></tr>
<tr><td>NVMe-IO-9</td><td>The device shall support the Compare and Write fused command pair.</td></tr>
<tr><td>NVMe-IO-10</td><td>For some models (see Section 12 Device Profiles), the device shall support the Write Uncorrectable command.</td></tr>
<tr><td>NVMe-IO-11</td><td>The Write Uncorrectable command shall support marking LBAs uncorrectable at a single LBA granularity regardless of the number of LBAs in the FTL indirection granularity.</td></tr>
<tr><td>NVMe-IO-12</td><td>The device shall not limit the number of LBAs that the host is able to specify in a Write Uncorrectable command beyond the minimum and maximum allowed by NVMe. The host shall be able to send a single LBA.</td></tr>
</table>

Requirement ID	Description
NVMe-IO-13	There shall be no limit on the total media capacity that can be marked uncorrectable by Write Uncorrectable commands.
NVMe-IO-14	Uncorrectable errors (e.g., read errors) that are a consequence of a prior Write Uncorrectable command shall not be counted in the Smart / Health Information (Log Identifier 02h) Media and Data Integrity Errors field.

4.6 Optional NVMe Feature Support

The device shall also support the following NVMe features:

Requirement ID	Description
NVMe-OPT-1	Obsolete. Duplicate of STD-LOG-7 (Telemetry Host-Initiated (Log Identifier 07h)) and STD-LOG-8 (Telemetry Controller-Initiated (Log Identifier 08h)).
NVMe-OPT-2	Timestamp (Feature Identifier 0Eh) shall be supported to align the devices internal logs.
NVMe-OPT-3	Background data collection by the device for either Host or Controller Initiated Telemetry shall not impact I/O latency or throughput.
NVMe-OPT-4	The device shall only clear the Timestamp Origin field to 000b in the Timestamp (Feature Identifier 0Eh) on a main power cycle or NVM Subsystem Reset (e.g., NSSR). The device shall not clear the Timestamp Origin field on a power cycle of only AUX power.
NVMe-OPT-5	The device shall never set the Synch field bit to 1b in the Timestamp (Feature Identifier 0Eh).
NVMe-OPT-6	The device shall only generate a Controller initiated Telemetry AER on an error condition and shall not generate a Controller initiated Telemetry AER for periodic logging.
NVMe-OPT-7	The device shall report its Indirection Unit (IU) size in the NPWG field in the Identify Namespace Data Structure.

4.7 Command Timeout

Requirement ID	Description
CTO-1	Admin and TCG Commands shall take no more than 10 seconds from submission to completion. CTO-1 does not apply to the time taken by background operations initiated by the Self-Test and Sanitize commands.
CTO-2	The only exceptions to CTO-1 shall be Format and the TCG commands Revert, RevertSP and GenKey.

Requirement ID	Description
CTO-3	Once TTR-2 is satisfied, an individual I/O command (QD1) shall take no more than 8 seconds from submission to completion. The device shall not have more than 7 I/Os take more than 2 seconds in one hour.
CTO-4	I/O command processing time shall not be a function of device capacity.
CTO-5	Device supplier shall disclose any I/O scenario that could violate the timeout requirements in CTO-1 through CTO-4.

4.8 Log Page Requirements

4.8.1 Standard Log Page Requirements

Requirement ID	Description
STD-LOG-1	Error Information (Log Identifier 01h) shall be supported.
STD-LOG-2	SMART / Health Information (Log Identifier 02h) shall be supported.
STD-LOG-3	Under no conditions shall the Percentage Used field in the SMART / Health Information (Log Identifier 02h) be reset.
STD-LOG-4	The Percentage Used field in the SMART / Health Information (Log Identifier 02h) shall be based on the average P/E cycle of the device. In addition, this field shall be based on the actual P/E cycle count of the media and not on the Power on Hours (POH) of the device.
STD-LOG-5	Firmware Slot Information (Log Identifier 03h) shall be supported.
STD-LOG-6	Commands Supported and Effects (Log Identifier 05h) shall be supported.
STD-LOG-7	Telemetry Host-Initiated (Log Identifier 07h) shall be supported.
STD-LOG-8	Telemetry Controller-Initiated (Log Identifier 08h) shall be supported.
STD-LOG-9	Persistent Event Log (Log Identifier 0Dh) shall be supported.
STD-LOG-10	The following Persistent Event Log types shall be supported:

Type	Event
01h	SMART / Health Log Snapshot
02h	Firmware Commit
03h	Timestamp Change
04h	Power-on or Reset
05h	NVM Subsystem Hardware Error
06h	Change Namespace
07h	Format NVM Start
08h	Format NVM Completion

Requirement ID	Description
	09h Sanitize Start 0Ah Sanitize Completion 0Ch Telemetry Log Created 0Dh Thermal Excursion

4.8.2 Telemetry Logging and Interface for Failure Analysis

The following applies to telemetry logging as the ability to quickly debug failures is required:

Requirement ID	Description
TEL-1	The device shall track the operational/event history and any critical parameters that can be used to debug issues.
TEL-2	The supplier shall provide a table that categorizes the reason identifiers that are a super set of the panic IDs in EREC-4 (Panic ID).
TEL-3	The device shall not lose debug data relevant to the current operating state due to any legal operation. Telemetry data including SMART / Health Information (Log Identifier 02h) and SMART / Health Information Extended (Log Identifier C0h) shall be persisted across power loss, all resets, Format command, Sanitize command, or device entering any protected mode (e.g., read only or panic mode).
TEL-4	The Reason Identifier field shall be the most recent failure identifier and shall not be cleared by a power cycle or reset.
TEL-5	The table below provides the specifications for the controller-initiated and the host-initiated log page "data areas": **Data Area** / **Purpose** / **Data Area Size** / **Latency Impact to IO** 1 / Periodic logging for monitoring trends/problems / Vendor-specific / < 1ms typical, 10ms max Implementation of Data areas 2 and 3 are optional and shall not be used for periodic logging.
TEL-6	Obsolete.
TEL-7	Obsolete.

4.8.3 DSSD Log Page Requirements

The table below defines the scope for all DSSD specific log pages:

Log Identifier	Scope	Log Page Name	Reference Section
C0h	NVM subsystem	SMART / Health Information Extended	4.8.5
C1h	NVM subsystem	Error Recovery	4.8.6
C2h	NVM subsystem	Firmware Activation History	4.8.7
C3h	Controller	Latency Monitor	4.8.9
C4h	NVM subsystem	Device Capabilities	4.8.10
C5h	NVM subsystem	Unsupported Requirements	4.8.11

KEY:
Namespace = The log page contains information about a specific namespace.
Controller = The log page contains information about the controller that is processing the command.
NVM subsystem = The log page contains information about the NVM subsystem.

4.8.4 SMART / Heath Information Extended (Log Identifier C0h) Requirements

Below are the requirements for the SMART / Health Information Extended (Log Identifier C0h):

Requirement ID	Description
SLOG-1	All values in the Vendor Log pages shall be persistent across power cycles unless otherwise specified.
SLOG-2	All counters shall be saturating counters (i.e., if the counter reaches the maximum allowable size, it stops incrementing and does NOT roll back to 0) unless otherwise specified.
SLOG-3	All values in logs shall be little endian format.
SLOG-4	A normalized counter, unless otherwise specified, shall be reported as the following: 100% shall represent the number at factory exit. 1% shall represent the minimum amount to be reliable. A value of 0% means the device shall no longer be considered reliable. 100% shall be represented as 64h.
SLOG-5	Devices shall support the attributes listed in Section 5.14.1.2 of the NVMe specification version 1.4b.
SLOG-6	A Read of either the SMART /Health Information (Log Identifier 02h) or SMART / Health Information Extended (Log Identifier C0h) shall not require an update of the SMART values other than the temperature values in SLOG-8. It shall be a simple read of the current data and shall not block IO.
SLOG-7	Unless otherwise specified, the device shall update these values in the background at least once every ten minutes.
SLOG-8	The composite and raw temperature sensor values shall be updated when the log page is accessed.

Requirement ID	Description
SLOG-9	All assert events and controller-initiated log captures will require an associated vendor-specific “Reason Identifier” that uniquely identifies the assert /controller condition.
SLOG-10	The device shall not lose any data in either the SMART / Health Information (Log Identifier 02h) or SMART / Health Information Extended (Log Identifier C0h) which is more than 10 minutes old including across power cycles/resets.
SLOG-11	The device shall not lose any back up energy source failure information or SMART / Health Information (Log Identifier 02h) critical warnings or SMART / Health Information Extended (Log Identifier C0h) critical warnings including across power cycles/resets.

4.8.5 SMART / Health Information Extended (Log Identifier C0h)

This vendor-specific log page, C0h shall be 512 bytes with the following functional requirements and field format:

Requirement ID	Byte Address	Field	# of Bytes	Field Description
SMART-1	15:0	Physical Media Units Written	16	Shall contain the number of bytes written to the media; this value includes both user and metadata written to the user and system areas. It shall be possible to use this attribute to calculate the Write Amplification Factor (WAF).
SMART-2	31:16	Physical Media Units Read	16	Shall contain the number of bytes read from the media from both the user and system areas.
SMART-3	39:32	Bad User NAND Blocks	8	The Raw count specifies the number of user NAND blocks that have been retired for any reason (e.g., program fails, erase fails or other events). The Normalized value is the percent of user spare blocks still available. The normalized value shall be set to 64h and the Raw count shall be cleared to zero on factory exit. It should be noted there are 2 bytes for normalized and 6 bytes for raw count. See SLOG-4 (normalized counter) definition above.

Byte Address	Field Description
39:38	Normalized value
37:32	Raw count

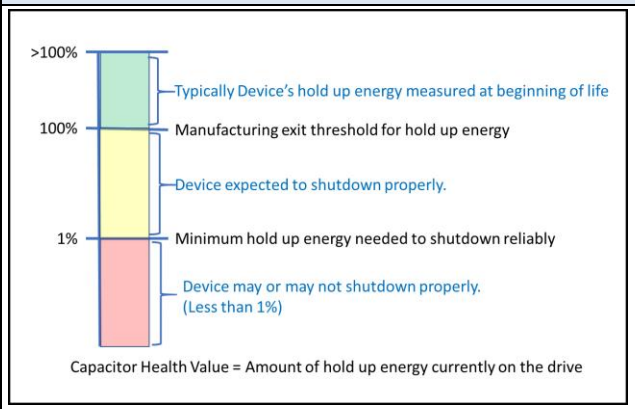
<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>SMART-4</td><td>47:40</td><td>Bad System NAND Blocks</td><td>8</td><td>The Raw count specifies the number of system NAND blocks that have been retired for any reason (e.g., program fails, erase fails or other events). The Normalized value is the percent of system spare blocks still available. The normalized value shall be set to 64h and the Raw count shall be cleared to zero on factory exit. It should be noted there are 2 bytes for normalized and 6 bytes for raw count. See SLOG-4 (normalized counter) definition above.<table><tr><th>Byte Address</th><th>Field Description</th></tr><tr><td>47:46</td><td>Normalized value</td></tr><tr><td>45:40</td><td>Raw count</td></tr></table>A value of FFFF_FFFF_FFFF_FFFFh indicates that the Bad User NAND block count field above represents all blocks on the device and Bad System NAND block count field is invalid.</td></tr>
<tr><td>SMART-5</td><td>55:48</td><td>XOR Recovery Count</td><td>8</td><td>Total number of times XOR was invoked to recover data in NAND. This shall cover all reads from NAND. Data recovery may have succeeded or failed. This shall be cleared to zero on factory exit.</td></tr>
<tr><td>SMART-6</td><td>63:56</td><td>Uncorrectable Read Error Count</td><td>8</td><td>Total count of NAND reads that were not correctable by read retries, all levels of ECC, or XOR. This shall be a count of the number of times data recovery fails and an uncorrectable read error is returned to the host.</td></tr>
<tr><td>SMART-7</td><td>71:64</td><td>Soft ECC Error Count</td><td>8</td><td>Total count of NAND reads that were not correctable by first level ECC and requires invoking an intermediate recovery. This shall cover all NAND read accesses. Data recovery may have succeeded or failed. If the device has more than one intermediate recovery level, then this counter only increments when intermediate recovery level 1 is invoked.</td></tr>
<tr><td>SMART-8</td><td>79:72</td><td>End to End Correction Counts</td><td>8</td><td>A count of the detected and corrected errors by the end-to-end error correction which includes DRAM, SRAM, or other storage element ECC/CRC protection mechanism (not NAND ECC).</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>All correctable errors shall result in a counter increase no matter what type of data the memory is protecting. All detected errors shall result in a counter increase unless the error is uncorrectable and occurred in the system region. In the latter case, the incomplete shutdown flag shall be flagged/incremented on the next power up. It should be noted there are 4 bytes for count of detected errors and 4 bytes for count of corrected errors.
<table>
<tr><th>Byte Address</th><th>Field Description</th></tr>
<tr><td>79:76</td><td>Corrected Errors</td></tr>
<tr><td>75:72</td><td>Detected Errors</td></tr>
</table></td></tr>
<tr><td>SMART-9</td><td>80</td><td>System Data % Used</td><td>1</td><td>A normalized cumulative count of the number of erase cycles per block since leaving the factory for the system (firmware and metadata) area. Starts at 0 and increments. 100 indicates that the estimated endurance has been consumed. Value may exceed 100 up to 255. This count shall increment regardless of what the backing media of the blocks are (e.g., SLC and TLC). If system data is split between media types, then this shall report the worst-case count so that the device wear out is clearly understood. This counter has a different behavior than the normalized counter definition in SLOG-4 (normalized counter), 100% (64h) represents the device many no longer function reliably as the max erase cycles has been hit.</td></tr>
<tr><td>SMART-10</td><td>87:81</td><td>Refresh Counts</td><td>7</td><td>This is a count of the number of blocks that have been re-allocated to maintain data integrity. This counter does not include creating free space due to garbage collection or block reallocation due to wear leveling.</td></tr>
<tr><td>SMART-11</td><td>95:88</td><td>User Data Erase Counts</td><td>8</td><td>The maximum and minimum erase counts across the user NAND blocks in the device. The host shall not be able to reset this counter. It should be noted there are 4 bytes for the maximum and 4 bytes for the minimum. The Minimum User</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>Data Erase Count shall not include bad blocks. If a block goes bad, any subsequent attempts to recover the block shall not increment the Maximum User Data Erase Count field.
<table>
<tr><th>Byte Address</th><th>Field Description</th></tr>
<tr><td>95:92</td><td>Minimum User Data Erase Count</td></tr>
<tr><td>91:88</td><td>Maximum User Data Erase Count</td></tr>
</table></td></tr>
<tr><td>SMART-12</td><td>97:96</td><td>Thermal Throttling Status and Count</td><td>2</td><td>The current status of thermal throttling (throttled or not throttled) and a count of the number of thermal throttling events. Note that there is 1 byte for the current status and 1 byte for the count. For devices that only have 1 throttle point only the first level throttle bit shall be set. This shall be cleared to zero on factory exit.
<table>
<tr><th>Byte Address</th><th>Field Description</th></tr>
<tr><td>97</td><td>Current Throttling Status

Current Status definition:
00h = unthrottled
01h = first level throttle
02h = 2nd level throttle
03h = 3rd level throttle
04h - FFh = Reserved</td></tr>
<tr><td>96</td><td>Number of thermal throttling events</td></tr>
</table></td></tr>
<tr><td>SMART-13</td><td>103:98</td><td>DSSD Specification Version</td><td>6</td><td>Version of the DSSD Specification that this device conforms to.
<table>
<tr><th>Byte Address</th><th>Field Description</th></tr>
<tr><td>103</td><td>Major Version Field. Shall be 02h.</td></tr>
<tr><td>102:101</td><td>Minor Version Field. Shall be 0000h.</td></tr>
<tr><td>100:99</td><td>Point Version Field. Shall be 0000h.</td></tr>
</table></td></tr>
</table>

Requirement ID	Byte Address	Field	# of Bytes	Field Description
				98 | Errata Version Field. Shall be 00h.
SMART-14	111:104	PCIe Correctable Error Count	8	Summation counter of all PCIe correctable errors (bad TLP, bad DLLP, receiver error, replay timeouts, replay rollovers). These counts shall only increment during run time. They shall not increment during training or power fail. This shall be cleared to zero on factory exit.
SMART-15	115:112	Incomplete Shutdowns	4	A count of the number of shutdowns that have occurred that did not completely flush all required user data and metadata to non-volatile memory for any reason. This shall be cleared to zero on factory exit.
SMART-16	119:116	Reserved	4	Shall be cleared to zero.
SMART-17	120	% Free Blocks	1	A normalized count of the number of blocks that are currently free (available) out of the total pool of spare (invalid) blocks. Free blocks means both blocks that have been erased and blocks that have all invalid data. Invalid blocks are blocks that are either marked invalid by device firmware or by the host (via de-allocate or overwrite). For example, if the total number of spare blocks is 100 and garbage collection has been able to reclaim (garbage collection and erase) 20 blocks, then this field reports 20%.
SMART-18	127:121	Reserved	7	Shall be cleared to zero.
SMART-19	129:128	Capacitor Health	2	This field is an indicator of the capacitor health and represents the capacitor holdup energy margin during operation. If no PLP protection is present a value of FFFFh shall be reported. 100% represents the passing hold up energy threshold when a device leaves manufacturing. Thus, a device will typically report greater than 100% in this field after leaving manufacturing at beginning of life. 1% is the minimum hold up energy required to conduct a proper shutdown reliably. A value of 0% may or may not result in a device failing to shutdown properly. This value shall never go negative. Zero is the minimum.

Requirement ID	Byte Address	Field	# of Bytes	Field Description
				
SMART-20	130	NVMe Errata Version	1	The device shall report the NVMe Errata revision (a, b, c, etc.) in ASCII that it conforms to. If there is no errata revision this field shall be cleared to zero.
SMART-30	135:131	Reserved	5	Shall be cleared to zero.
SMART-21	143:136	Unaligned I/O	8	This is a count of the number of write IOs performed by the device that are not aligned to the indirection unit size (IU) of the device. Alignment indicates only the start of each IO. The length does not affect this count. This counter shall reset on power cycle. This counter shall not wrap. This shall be cleared to zero on factory exit.
SMART-22	151:144	Security Version Number	8	This is the Security Version Number of the currently running firmware image. The supplier increments this number any time a firmware includes a fix for a security issue or critical firmware fix that customer agrees rollback prevention is required.
SMART-23	159:152	Total NUSE	8	Total Namespace Utilization. For a device with a single Namespace, this shall be a copy of the Namespace Utilization field defined in the Identify Namespace Data Structure bytes 23:16. For a device with multiple Namespaces, this shall reflect the total utilization based on all the Namespaces.
SMART-24	175:160	PLP Start Count	16	This is a count of the number of times the device has initiated its power loss protection process due to supply voltage drop. This counter shall be incremented on the initial detection of the

Requirement ID	Byte Address	Field	# of Bytes	Field Description
				power loss condition. This does not include PLP health check operations.
SMART-25	191:176	Endurance Estimate	16	This field is an estimate of the total number of data bytes that may be written to the device over its lifetime assuming a write amplification of 1. (i.e., no increase in the number of write operations performed by the device beyond the number of write operations requested by a host). This value shall be equivalent to the Endurance Estimate field in the Endurance Group Log (Log Identifier 09h).
SMART-29	199:192	PCIe Link Retraining Count	8	This is a count of the number of PCIe Link Retraining events. This count shall only increment during run time. It shall not increment during training or power fail. This shall be cleared to zero on factory exit.
SMART-31	207:200	Power State Change Count	8	Summation counter of the number of NVMe Power State changes whether host or device initiated, including NVMe Power State changes caused by DSSD Power State changes (see Section 4.8.10.1 DSSD Power State Requirements). This count shall only increment during run time. This shall be cleared to zero on factory exit.
SMART-26	493:208	Reserved	286	Shall be cleared to zero.
SMART-27	495:494	Log Page Version	2	This indicates the version of the mapping this log page uses. Shall be set to 0003h.
SMART-28	511:496	Log Page GUID	16	Shall be set to AFD514C97C6F4F9CA4f2BFEA2810AFC5h.

4.8.6 Error Recovery (Log Identifier C1h)

Below are the requirements for the Error Recovery (Log Identifier C1h) AEN requirements:

Requirement ID	Description
EREC-AEN-1	If the device encounters a Panic condition, it shall complete an outstanding Asynchronous Event Request (AER) command with an Asynchronous Event Notification (AEN) with Completion Queue Entry Dword 0 set as follows: • The Log Page Identifier field shall be set to C1h. • The Asynchronous Event Information field shall be cleared to zero.

Requirement ID	Description
	• The Asynchronous Event Type field shall be set to 111b (Vendor Specific).

This vendor-specific log page, C1h shall be 512 bytes with the following functional requirements and field format:

Requirement ID	Byte Address	Field	# of Bytes	Field Description
EREC -1	1:0	Panic Reset Wait Time	2	The amount of time the host should wait for the device panic workflow to complete in msec.
EREC-2	2	Panic Reset Action	1	Bit field indicating potential reset actions that may need to be taken. If no reset action needed, do not set any of the bits. More than 1 bit can be set, and it is up to the host to decide the sequence of action(s) to take. Use Bit 0 if possible. If Bit 0 is not possible use Bit 1, etc. **Bit** / **Description** 7:6 – Reserved. Shall be cleared to zero. 5 – PCIe Conventional Hot Reset. 4 – Main Power Cycle. 3 – PERST#. 2 – PCIe Function Level Reset. 1 – NVM Subsystem Reset. 0 – NVMe Controller Reset.
EREC-3	3	Device Recovery Action 1	1	The recovery action to take for handling a device panic condition. Value is dependent on the panic condition. Use 01h if possible. **Bit** / **Description** 7:6 – Reserved. Shall be cleared to zero. 5 – Sanitize Required. 4 – Device Replacement Required. 3 – Vendor Analysis Required. 2 – Vendor Specific Command Required. 1 – Format NVM Required (Any SES value, any supported combination of other parameters - PIL, PI, MSET, LBAF). 0 – No Action Required.

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>EREC-4</td><td>11:4</td><td>Panic ID</td><td>8</td><td>ID to identify the panic condition encountered. A Zero value indicates no panic. Value is dependent on the panic condition.
The following Panic ID values are reserved for Host defined fault codes for known panic conditions:
○ 00000000_00000000h – 00000000_0000FFFFh
<table><tr><th>Byte</th><th>Field Description</th></tr><tr><td>11:4</td><td>Panic ID definition:
• 00000000_00000001h – Panic caused by flush failures or data loss during power loss handling.</td></tr></table></td></tr>
<tr><td>EREC-5</td><td>15:12</td><td>Device Capabilities</td><td>4</td><td>Field to indicate device capabilities.
<table><tr><th>Bit</th><th>Description</th></tr><tr><td>31:2</td><td>Reserved. Shall be cleared to zero.</td></tr><tr><td>1</td><td>Panic CFS Supported: If set, indicates device supports using CFS to notify host of a panic condition*.</td></tr><tr><td>0</td><td>Panic AEN Supported: If set, indicates device supports using AEN to notify host of a panic condition*.</td></tr></table>*Note: It is valid for a device to indicate support for both Panic AEN Supported and Panic Controller Fatal Status Supported. If the device supports both, the device shall only use one of the panic notification mechanisms when reporting a given panic event.</td></tr>
<tr><td>EREC-6</td><td>16</td><td>Vendor Specific Recovery Opcode</td><td>1</td><td>Vendor specific command opcode to recover device from panic condition. Only valid when Device Recovery Action field value is 02h. When Device Recovery Action 1 field value is not 02h, this field shall be cleared to zero.</td></tr>
<tr><td>EREC-7</td><td>19:17</td><td>Reserved</td><td>3</td><td>Shall be cleared to zero.</td></tr>
<tr><td>EREC-8</td><td>23:20</td><td>Vendor Specific Command CDW12</td><td>4</td><td>CDW12 value for the Vendor Specific command to recover device from panic condition. Only valid when Device Recovery Action field value is</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>02h. When Device Recovery Action 1 field value is not 02h, this field shall be cleared to zero.</td></tr>
<tr><td>EREC-9</td><td>27:24</td><td>Vendor Specific Command CDW13</td><td>4</td><td>CDW13 value for the Vendor Specific command to recover device from panic condition. Only valid when Device Recovery Action 1 field value is 02h. When Device Recovery Action 1 field value is not 02h, this field shall be cleared to zero.</td></tr>
<tr><td>EREC-13</td><td>28</td><td>Vendor Specific Command Timeout</td><td>1</td><td>The amount of time the host should wait for the device to complete the recovery command in seconds.</td></tr>
<tr><td>EREC-14</td><td>29</td><td>Device Recovery Action 2</td><td>1</td><td>Bit field indicating potential post reset actions that may need to be taken. If no reset action needed, do not set any of the bits. More than 1 bit can be set, and it is up to the host to decide the sequence of action(s) to take. Use Bit 0 if possible. If Bit 0 is not possible use Bit 1, etc.
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>7:6</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>5</td><td>PCIe Conventional Hot Reset.</td></tr>
<tr><td>4</td><td>Main Power Cycle.</td></tr>
<tr><td>3</td><td>PERST#.</td></tr>
<tr><td>2</td><td>PCIe Function Level Reset.</td></tr>
<tr><td>1</td><td>NVM Subsystem Reset.</td></tr>
<tr><td>0</td><td>NVMe Controller Reset.</td></tr>
</table></td></tr>
<tr><td>EREC-15</td><td>30</td><td>Device Recovery Action 2 Timeout</td><td>1</td><td>The amount of time the host should wait for the device to complete the Device Recovery Action 2 in seconds (overrides CAP.TO value).</td></tr>
<tr><td>EREC-10</td><td>493:31</td><td>Reserved</td><td>463</td><td>Shall be cleared to zero.</td></tr>
<tr><td>EREC-11</td><td>495:494</td><td>Log Page Version</td><td>2</td><td>This indicates the version of the mapping this log page uses. Shall be set to 0002h.</td></tr>
<tr><td>EREC-12</td><td>511:496</td><td>Log Page GUID</td><td>16</td><td>Shall be set to 5A1983BA3DFD4DABAE3430FE2131D944h.</td></tr>
</table>

4.8.7 Firmware Activation History (Log Identifier C2h) Requirements

This defines the requirements for recording the Firmware Activation History.

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td>FWHST-LOG-1</td><td>Lists the last twenty firmware images that were activated (not downloaded) on the device. This is a circular buffer where the 21st entry is placed in entry 0 (byte offset 8 decimal).</td></tr>
<tr><td>FWHST-LOG-2</td><td>When the device is first shipped from the factory, there are no entries recorded.</td></tr>
<tr><td>FWHST-LOG-3</td><td>A Firmware Activation History Entry shall be recorded as shown in the table below whenever a new firmware is activated by the device. Firmware downloads shall not generate an entry.
<table>
<tr><th>Commit Action</th><th>Response</th></tr>
<tr><td>000b</td><td>No Firmware Activation History Entry is recorded.</td></tr>
<tr><td>001b</td><td>A Firmware Activation History Entry is recorded after the next Controller Level Reset.</td></tr>
<tr><td>010b</td><td>A Firmware Activation History Entry is recorded after the next Controller Level Reset.</td></tr>
<tr><td>011b</td><td>A Firmware Activation History Entry is recorded immediately.</td></tr>
</table></td></tr>
<tr><td>FWHST-LOG-4</td><td>Redundant activation events shall not generate a new entry to prevent the scrolling out of useful information. An entry shall be considered redundant if it meets ALL the criteria below:
1. Timestamp is within 1 minute from the last RECORDED entry; and
2. Power cycle count is the same; and
3. Current firmware is the same; and
4. New firmware activated is the same; and
5. Slot number is the same; and
6. Commit Action Type is the same; and
7. The Result field has not changed.</td></tr>
<tr><td>FWHST-LOG-5</td><td>Firmware Activation History’s log page format shall follow the requirements below.</td></tr>
</table>

4.8.7.1 Firmware Activation History (Log Identifier C2h)

This vendor-specific log page, C2h shall be 4096 bytes with the following functional requirements and field format:

Requirement ID	Byte Address	Field	# of Bytes	Field Description
FAHL-1	0	Log Identifier	1	This field shall be set to C2h.
FAHL-2	3:1	Reserved	3	Shall be cleared to zero.

Requirement ID	Byte Address	Field	# of Bytes	Field Description
FAHL-3	7:4	Valid Firmware Activation History Entries	4	Contains the number of event entries in the log that are valid. Starts at 0 from the factory or after a Clear Firmware Update History (Feature Identifier C1h) Set Feature (see Section 4.12.4 Clear Firmware Update History (Feature Identifier C1h) Set Feature). Increments on each new log entry up to the maximum of 20 entries (see FWHST-LOG-4 (Redundant activation events shall not generate a new entry)).
FAHL-4	71:8	Firmware Activation History Entry 0	64	This field contains the first firmware activation entry.
	…	…	…	…
	1287:1224	Firmware Activation History Entry 19	64	This field contains the last firmware activation entry.
FAHL-5	4077:1288	Reserved	2790	Shall be cleared to zero.
FAHL-6	4079:4078	Log Page Version	2	This indicates the version of the mapping this log page uses. Shall be set to 0001h.
FAHL-7	4095:4080	Log Page GUID	16	Shall be set to D11CF3AC8AB24DE2A3F6DAB4769A796Dh.

4.8.7.2 Firmware Activation History Entry Format

This defines the History Entry format for recording Firmware Activation History events.

Requirement ID	Byte Address	Field	# of Bytes	Field Description
FAHE-1	0	Entry Version Number	1	Indicates the version of this entry format used in the device. Shall be set to 01h.
FAHE-2	1	Entry Length (EL)	1	This field indicates the length in bytes of the entry log event data. Shall be set to 40h.
FAHE-3	3:2	Reserved	2	Shall be cleared to zero.
FAHE-4	5:4	Firmware Activation Count	2	This field shall increment every time a firmware activation is attempted regardless of the result. This value shall be cleared to zero when the device is shipped from the factory. This field

Requirement ID	Byte Address	Field	# of Bytes	Field Description
				shall be a saturating counter. This field is not affected by a Clear Firmware History (Feature Identifier C1h) Set Feature and continues counting from the next integer on the next Firmware Activation.
FAHE-5	13:6	Timestamp	8	This field shall indicate the Timestamp of when the new firmware activation is completed, and the device is running with the new firmware. The format of this field shall be as defined in Section 5.21.1.14 Timestamp (Feature Identifier 0Eh) of the NVMe 1.4b specification.
FAHE-6	21:14	Reserved	8	Shall be cleared to zero.
FAHE-7	29:22	Power Cycle Count	8	This field shall indicate the power cycle count in which the firmware activation occurred.
FAHE-8	37:30	Previous Firmware	8	This field shall indicate the previous firmware version running on the device before this firmware activation took place. The format of this field shall be as defined in field Firmware Revision (FR) Section 5.15.2.2 Identify Controller Data Structure of the NVMe 1.4b specification.
FAHE-9	45:38	New Firmware Activated	8	This field shall indicate the activated firmware version that is running on the device after the firmware activation took place. If firmware activation failed, then this field shall display the same information as FAHE-8. The format of this field shall be as defined in field Firmware Revision (FR) Section 5.15.2.2 Identify Controller Data Structure of the NVMe 1.4b specification.
FAHE-10	46	Slot Number	1	This field shall indicate the slot that the activated firmware is in.
FAHE-11	47	Commit Action Type	1	This field shall indicate the Commit action type associated with the firmware activation event.
FAHE-12	49:48	Result	2	This field shall indicate the results of the firmware activation event. A value of 0000h shall represent the firmware commit was successful. A non-zero value shall represent the firmware commit was unsuccessful and the value represents the status code associated with the failure.

Requirement ID	Byte Address	Field	# of Bytes	Field Description
FAHE-13	63:50	Reserved	14	Shall be cleared to zero.

4.8.8 Latency Monitor Log and Feature Set Requirements

The following are requirements for the Latency Monitor Log and Feature Set. For more information about the Latency Monitoring Feature Set (see Appendix C Latency Monitoring Feature Set Theory of Operation).

Requirement ID	Description
LMLOG-1	All values in the Latency Monitor Log Page shall be persistent across power cycles and resets unless otherwise specified.
LMLOG-2	All counters shall be saturating counters (i.e., if the counter reaches the maximum allowable size, it stops incrementing and does NOT roll back to 0).
LMLOG-3	All values in the Latency Monitor Log shall be little endian format.
LMLOG-4	A read of the Latency Monitor Log shall be a simple read of the active data and shall not block IO.
LMLOG-5	Data in the Latency Monitor Log which is read by the host shall be no more than 10 minutes old.
LMLOG-6	When configuring the Latency Monitoring Feature with Set Features the Active Buckets and Static Buckets shall be reset.
LMLOG-7	When powering on, the counters shall run based on the previously configured values. If the values have never been configured, they shall run based on the default values.
LMLOG-8	The Latency Stamp shall be based on the NVMe timestamp command if it has been received. If the NVMe timestamp command has not been received it shall be based on the device power on hours.
LMLOG-9	When the device provides a Latency Stamp of latency outliers, the Latency Stamp shall be based on command completion.
LMLOG-10	When generating a Latency Monitoring Log, the latency shall be no greater than the latency associated with generating a Telemetry log.
LMLOG-11	When configuring this feature, the thresholds shall always be configured such that Active Threshold A < Active Threshold B < Active Threshold C < Active Threshold D. If the host attempts to configure the device in such a way that violates the above rules, the device shall return Invalid Field in Command.
LMLOG-12	The device shall support the Latency Monitoring Feature Set.
LMLOG-13	The Latency Monitor Log page shall be 512 bytes.
LMLOG-14	Executing a Set Feature command for Latency Monitor (Feature Identifier C5h) shall reset all the contents in the Active/Static Buckets.

4.8.9 Latency Monitor (Log Identifier C3h)

This vendor-specific log page, C3h shall be 512 bytes with the following functional requirements and field format (see Appendix C Latency Monitoring Feature Set Theory of Operation for additional details):

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>LMDATA-1</td><td>0</td><td>Latency Monitor Feature Status</td><td>1</td><td>
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>7:3</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>2</td><td>Active Measured Latency Supported.

When set to 1b the Active Measured Latency field is supported and shall be populated based on the Active Latency Configuration settings. When cleared to 0b the Active Measured Latency is not supported.</td></tr>
<tr><td>1</td><td>Active Latency Configuration/Active Latency Mode 1 Supported.

When set to 1b the device shall support the Active Latency Configuration with the Active Latency Mode = 1b. When cleared to 0b the device does not support Active Latency Mode =1b.</td></tr>
<tr><td>0</td><td>Latency Monitoring Feature Enabled

This is a global feature enable. When set to 1b the Latency Monitoring Feature for this device is enabled. When cleared to 0b all features in the Latency Monitoring Log page are disabled for this device and can be ignored. When cleared to 0b the other fields in this log page are not power loss safe and may be lost.</td></tr>
</table>
The default value shall be 07h.
</td></tr>
<tr><td>LMDATA-2</td><td>1</td><td>Reserved</td><td>1</td><td>Shall be cleared to zero.</td></tr>
</table>

Requirement ID	Byte Address	Field	# of Bytes	Field Description
LMDATA-3	3:2	Active Bucket Timer	2	The Active Bucket Timer is in 5-minute increments. Thus, a value of 0001h is 5 minutes and a value of 0002h is 10 minutes. This represents the amount of time the Active Buckets have been accumulating data. The Active Bucket Timer will saturate at FFFFh. When the Active Bucket Timer reaches the Active Bucket Timer Threshold then the data specified in Appendix C – Latency Monitoring Feature Set Theory of Operation is moved into the Static Buckets, the Active Bucket Timer is cleared to 0000h and restarts counting. If the Active Bucket Timer is running and there is a power cycle the Active Bucket Timer value from before the power cycle shall be restored into the Active Bucket Timer and the Active Bucket Timer shall continue when the device is powered on.
LMDATA-4	5:4	Active Bucket Timer Threshold	2	Active Bucket Timer Threshold is the threshold used to compare with the Active Bucket Timer. When cleared to 0000h this threshold is not used, the Active Bucket Timer will saturate, and the Static Buckets will not be loaded. This threshold is in 5-minute increments. The factory default value of the Active Bucket Timer Threshold shall be set to 07E0h.
LMDATA-5	6	Active Threshold A	1	This defines Active Threshold A. This is in 5ms increments. A value of 00h represents 5ms. A value of FFh represents 1.280 seconds. The factory default is 05h.
LMDATA-6	7	Active Threshold B	1	This is in 5ms increments. The factory default is 13h.
LMDATA-7	8	Active Threshold C	1	This is in 5ms increments. The factory default is 1Eh.
LMDATA-8	9	Active Threshold D	1	This is in 5ms increments. The factory default is 2Eh.
LMDATA-9	11:10	Active Latency	2	This configures how both the Active Latency Stamp, and the Active Measured Latency Fields

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td>Configuration</td><td></td><td>are updated on a per I/O command (Read, Write, Deallocate) counter basis.

When the Active Latency Mode is cleared to 0b the Active Latency Stamp and the Active Measured Latency will trigger and update the first time the associated command counter increments. Once this trigger happens the fields shall not be updated until the Active Latency Stamp and Active Measured Latency fields are reset based on the Active Bucket Timer expiring.

When the Active Latency Mode is set to 1b the Active Latency Stamp and the Active Measured Latency fields shall update to show the largest measured latency based on the associated command counter.
<table>
<tr><th>Bit</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>15:12</td><td>N/A</td><td>N/A</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>11</td><td>3</td><td>De-allocate/TRIM</td><td>Active Latency Mode [11]</td></tr>
<tr><td>10</td><td>3</td><td>Write</td><td>Active Latency Mode [10]</td></tr>
<tr><td>9</td><td>3</td><td>Read</td><td>Active Latency Mode [9]</td></tr>
<tr><td>8</td><td>2</td><td>De-allocate/TRIM</td><td>Active Latency Mode [8]</td></tr>
<tr><td>7</td><td>2</td><td>Write</td><td>Active Latency Mode [7]</td></tr>
<tr><td>6</td><td>2</td><td>Read</td><td>Active Latency Mode [6]</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>5</td><td>1</td><td>De-allocate/TRIM</td><td>Active Latency Mode [5]</td></tr>
<tr><td>4</td><td>1</td><td>Write</td><td>Active Latency Mode [4]</td></tr>
<tr><td>3</td><td>1</td><td>Read</td><td>Active Latency Mode [3]</td></tr>
<tr><td>2</td><td>0</td><td>De-allocate/TRIM</td><td>Active Latency Mode [2]</td></tr>
<tr><td>1</td><td>0</td><td>Write</td><td>Active Latency Mode [1]</td></tr>
<tr><td>0</td><td>0</td><td>Read</td><td>Active Latency Mode [0]</td></tr>
</table>
The default value shall be 0FFFh.
</td></tr>
<tr><td>LMDATA-10</td><td>12</td><td>Active Latency Minimum Window</td><td>1</td><td>This is the minimum number of 100 milliseconds increments between Latency Events for a single Active Latency Stamp and Active Measured Latency. When cleared to 00h this feature is disabled. This count is in 100 millisecond increments, thus a value of 01h is 100 milliseconds and 02h is 200 milliseconds. Once a Latency Stamp/Measured Latency is updated if the Active Latency Minimum Window time has not expired and an event that is configurated to generate a Latency Stamp/Measured Latency occurs the Latency Stamp/Measured Latency will not be recorded. The default value of this field is 0Ah.</td></tr>
<tr><td>LMDATA-11</td><td>31:13</td><td>Reserved</td><td>19</td><td>Shall be cleared to zero.</td></tr>
<tr><td>LMDATA-12</td><td>47:32</td><td>Active Bucket Counter 0</td><td>16</td><td>
<table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>47:44</td><td>Read Command Counter.</td></tr>
<tr><td>43:40</td><td>Write Command Counter.</td></tr>
<tr><td>39:36</td><td>De-Allocate/TRIM Command Counter.</td></tr>
<tr><td>35:32</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>LMDATA-13</td><td>63:48</td><td>Active Bucket Counter 1</td><td>16</td><td>
<table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>63:60</td><td>Read Command Counter.</td></tr>
<tr><td>59:56</td><td>Write Command Counter.</td></tr>
<tr><td>55:52</td><td>De-Allocate/TRIM Command Counter.</td></tr>
<tr><td>51:48</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-14</td><td>79:64</td><td>Active Bucket Counter 2</td><td>16</td><td>
<table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>79:76</td><td>Read Command Counter.</td></tr>
<tr><td>75:72</td><td>Write Command Counter.</td></tr>
<tr><td>71:68</td><td>De-Allocate/TRIM Command Counter.</td></tr>
<tr><td>67:64</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-15</td><td>95:80</td><td>Active Bucket Counter 3</td><td>16</td><td>
<table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>95:92</td><td>Read Command Counter.</td></tr>
<tr><td>91:88</td><td>Write Command Counter.</td></tr>
<tr><td>87:84</td><td>De-Allocate/TRIM Command Counter.</td></tr>
<tr><td>83:80</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-16</td><td>191:96</td><td>Active Latency Stamp</td><td>96</td><td>This field contains a Timestamp for when a latency event occurred for each counter. A value of FFFFFFFF_FFFFFFFFh means the Latency Stamp is not valid. The Active Latency Stamp uses the data format for Timestamp as defined in NVMe.
<table>
<tr><th>Byte Address</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>191:184</td><td>0</td><td>Read</td><td>Active Latency Stamp 0</td></tr>
<tr><td>183:176</td><td>0</td><td>Write</td><td>Active Latency Stamp 1</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>175:168</td><td>0</td><td>De-allocate/ TRIM</td><td>Active Latency Stamp 2</td></tr>
<tr><td>167:160</td><td>1</td><td>Read</td><td>Active Latency Stamp 3</td></tr>
<tr><td>159:152</td><td>1</td><td>Write</td><td>Active Latency Stamp 4</td></tr>
<tr><td>151:144</td><td>1</td><td>De-allocate/ TRIM</td><td>Active Latency Stamp 5</td></tr>
<tr><td>143:136</td><td>2</td><td>Read</td><td>Active Latency Stamp 6</td></tr>
<tr><td>135:128</td><td>2</td><td>Write</td><td>Active Latency Stamp 7</td></tr>
<tr><td>127:120</td><td>2</td><td>De-allocate/ TRIM</td><td>Active Latency Stamp 8</td></tr>
<tr><td>119:112</td><td>3</td><td>Read</td><td>Active Latency Stamp 9</td></tr>
<tr><td>111:104</td><td>3</td><td>Write</td><td>Active Latency Stamp 10</td></tr>
<tr><td>103:96</td><td>3</td><td>De-allocate/ TRIM</td><td>Active Latency Stamp 11</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-17</td><td>215:192</td><td>Active Measured Latency</td><td>24</td><td>This is the measured latency that caused the counter to increment. A value of 0000h means this field is invalid. A value of 0001h represents 1 millisecond. A value of 0002h represents 2 milliseconds.
<table>
<tr><th>Byte Address</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>215:214</td><td>0</td><td>Read</td><td>Active Measured Latency 0</td></tr>
<tr><td>213:212</td><td>0</td><td>Write</td><td>Active Measured Latency 1</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>211:210</td><td>0</td><td>De-allocate/ TRIM</td><td>Active Measured Latency 2</td></tr>
<tr><td>209:208</td><td>1</td><td>Read</td><td>Active Measured Latency 3</td></tr>
<tr><td>207:206</td><td>1</td><td>Write</td><td>Active Measured Latency 4</td></tr>
<tr><td>205:204</td><td>1</td><td>De-allocate/ TRIM</td><td>Active Measured Latency 5</td></tr>
<tr><td>203:202</td><td>2</td><td>Read</td><td>Active Measured Latency 6</td></tr>
<tr><td>201:200</td><td>2</td><td>Write</td><td>Active Measured Latency 7</td></tr>
<tr><td>199:198</td><td>2</td><td>De-allocate/ TRIM</td><td>Active Measured Latency 8</td></tr>
<tr><td>197:196</td><td>3</td><td>Read</td><td>Active Measured Latency 9</td></tr>
<tr><td>195:194</td><td>3</td><td>Write</td><td>Active Measured Latency 10</td></tr>
<tr><td>193:192</td><td>3</td><td>De-allocate/ TRIM</td><td>Active Measured Latency 11</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-18</td><td>217:216</td><td>Active Latency Stamp Units</td><td>2</td><td>When bit is set to 1b the Active Latency Stamp was based on receiving the NVMe Timestamp. When bit is cleared to 0b the Active Latency Stamp was based on power on hours since the NVMe Timestamp was not received.
<table>
<tr><th>Bit</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>15:12</td><td>N/A</td><td>N/A</td><td>Reserved. Shall be</td></tr>
</table>
</td></tr>
</table>

Requirement ID	Byte Address	Field	# of Bytes	Field Description
				<table><tr><td></td><td></td><td></td><td>cleared to zero.</td></tr><tr><td>11</td><td>3</td><td>De-allocate/TRIM</td><td>Active Latency Stamp Unit [11]</td></tr><tr><td>10</td><td>3</td><td>Write</td><td>Active Latency Stamp Unit [10]</td></tr><tr><td>9</td><td>3</td><td>Read</td><td>Active Latency Stamp Unit [9]</td></tr><tr><td>8</td><td>2</td><td>De-allocate/TRIM</td><td>Active Latency Stamp Unit [8]</td></tr><tr><td>7</td><td>2</td><td>Write</td><td>Active Latency Stamp Unit [7]</td></tr><tr><td>6</td><td>2</td><td>Read</td><td>Active Latency Stamp Unit [6]</td></tr><tr><td>5</td><td>1</td><td>De-allocate/TRIM</td><td>Active Latency Stamp Unit [5]</td></tr><tr><td>4</td><td>1</td><td>Write</td><td>Active Latency Stamp Unit [4]</td></tr><tr><td>3</td><td>1</td><td>Read</td><td>Active Latency Stamp Unit [3]</td></tr><tr><td>2</td><td>0</td><td>De-allocate/TRIM</td><td>Active Latency Stamp Unit [2]</td></tr><tr><td>1</td><td>0</td><td>Write</td><td>Active Latency Stamp Unit [1]</td></tr><tr><td>0</td><td>0</td><td>Read</td><td>Active Latency Stamp Unit [0]</td></tr></table>
LMDATA-19	239:218	Reserved	22	Shall be cleared to zero.
LMDATA-20	255:240	Static Bucket Counter 0	16	This is a snapshot of the Active Bucket Counter 0 which is moved to this field when the Active Bucket Timer equals the Active Bucket Timer Threshold.

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td><table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>255:252</td><td>Read Command Counter.</td></tr>
<tr><td>251:248</td><td>Write Command Counter.</td></tr>
<tr><td>247:244</td><td>De-Allocate/TRIM Command Counter.</td></tr>
<tr><td>243:240</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table></td></tr>
<tr><td>LMDATA-21</td><td>271:256</td><td>Static Bucket Counter 1</td><td>16</td><td>This is a snapshot of the Active Bucket Counter 1 which is moved to this field when the Active Bucket Timer equals the Active Bucket Timer Threshold.<table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>271:268</td><td>Read Command Counter.</td></tr>
<tr><td>267:264</td><td>Write Command Counter.</td></tr>
<tr><td>263:260</td><td>De-Allocate/TRIM Command Counter.</td></tr>
<tr><td>259:256</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table></td></tr>
<tr><td>LMDATA-22</td><td>287:272</td><td>Static Bucket Counter 2</td><td>16</td><td>This is a snapshot of the Active Bucket Counter 2 which is moved to this field when the Active Bucket Timer equals the Active Bucket Timer Threshold.<table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>287:284</td><td>Read Command Counter.</td></tr>
<tr><td>283:280</td><td>Write Command Counter.</td></tr>
<tr><td>279:276</td><td>De-Allocate/TRIM Command Counter.</td></tr>
<tr><td>275:272</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table></td></tr>
<tr><td>LMDATA-23</td><td>303:288</td><td>Static Bucket Counter 3</td><td>16</td><td>This is a snapshot of the Active Bucket Counter 3 which is moved to this field when the Active Bucket Timer equals the Active Bucket Timer Threshold.</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>303:300</td><td>Read Command Counter.</td></tr>
<tr><td>299:296</td><td>Write Command Counter.</td></tr>
<tr><td>295:292</td><td>De-Allocate/TRIM Command Counter.</td></tr>
<tr><td>291:288</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-24</td><td>399:304</td><td>Static Latency Stamp</td><td>96</td><td>This is a snapshot of the Active Latency Stamp which is moved to this field when the Active Bucket Timer equals the Active Bucket Timer Threshold.

This field contains a timestamp for when a latency event occurred for each counter. A value of FFFFFFFF_FFFFFFFFh means the Latency Stamp is not valid. The Static Latency Stamp uses the Timestamp data format as defined in NVMe.
<table>
<tr><th>Byte Address</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>399:392</td><td>0</td><td>Read</td><td>Static Latency Stamp 0</td></tr>
<tr><td>391:384</td><td>0</td><td>Write</td><td>Static Latency Stamp 1</td></tr>
<tr><td>383:376</td><td>0</td><td>De-allocate/ TRIM</td><td>Static Latency Stamp 2</td></tr>
<tr><td>375:368</td><td>1</td><td>Read</td><td>Static Latency Stamp 3</td></tr>
<tr><td>367:360</td><td>1</td><td>Write</td><td>Static Latency Stamp 4</td></tr>
<tr><td>359:352</td><td>1</td><td>De-allocate/ TRIM</td><td>Static Latency Stamp 5</td></tr>
<tr><td>351:344</td><td>2</td><td>Read</td><td>Static Latency Stamp 6</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>343:336</td><td>2</td><td>Write</td><td>Static Latency Stamp 7</td></tr>
<tr><td>335:328</td><td>2</td><td>De-allocate/ TRIM</td><td>Static Latency Stamp 8</td></tr>
<tr><td>327:320</td><td>3</td><td>Read</td><td>Static Latency Stamp 9</td></tr>
<tr><td>319:312</td><td>3</td><td>Write</td><td>Static Latency Stamp 10</td></tr>
<tr><td>311:304</td><td>3</td><td>De-allocate/ TRIM</td><td>Static Latency Stamp 11</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-25</td><td>423:400</td><td>Static Measured Latency</td><td>24</td><td>This is a snapshot of the Active Measured Latency which is moved to this field when the Active Bucket Timer equals the Active Bucket Timer Threshold.

A value of 0000h means this field is invalid. A value of 0001h represents 1 millisecond. A value of 0002h represents 2 milliseconds.
<table>
<tr><th>Byte Address</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>423:422</td><td>0</td><td>Read</td><td>Static Measured Latency 0</td></tr>
<tr><td>421:420</td><td>0</td><td>Write</td><td>Static Measured Latency 1</td></tr>
<tr><td>419:418</td><td>0</td><td>De-allocate/ TRIM</td><td>Static Measured Latency 2</td></tr>
<tr><td>417:416</td><td>1</td><td>Read</td><td>Static Measured Latency 3</td></tr>
<tr><td>415:414</td><td>1</td><td>Write</td><td>Static Measured Latency 4</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>413:412</td><td>1</td><td>De-allocate/ TRIM</td><td>Static Measured Latency 5</td></tr>
<tr><td>411:410</td><td>2</td><td>Read</td><td>Static Measured Latency 6</td></tr>
<tr><td>409:408</td><td>2</td><td>Write</td><td>Static Measured Latency 7</td></tr>
<tr><td>407:406</td><td>2</td><td>De-allocate/ TRIM</td><td>Static Measured Latency 8</td></tr>
<tr><td>405:404</td><td>3</td><td>Read</td><td>Static Measured Latency 9</td></tr>
<tr><td>403:402</td><td>3</td><td>Write</td><td>Static Measured Latency 10</td></tr>
<tr><td>401:400</td><td>3</td><td>De-allocate/ TRIM</td><td>Static Measured Latency 11</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-26</td><td>425:424</td><td>Static Latency Stamp Units</td><td>2</td><td>This is a snapshot of the Active Latency Stamp Units which is moved to this field when the Active Bucket Timer equals the Active Bucket Timer Threshold.

When bit is set to 1b the Static Latency Stamp was based on receiving the NVMe Timestamp and the offset from this. When bit is cleared to 0b the Static Latency Stamp was based on power on hours since the NVMe Timestamp was not received.
<table>
<tr><th>Bit</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>15:12</td><td>N/A</td><td>N/A</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><td>11</td><td>3</td><td>De-allocate/TRIM</td><td>Static Latency Stamp Unit [11]</td></tr>
<tr><td>10</td><td>3</td><td>Write</td><td>Static Latency Stamp Unit [10]</td></tr>
<tr><td>9</td><td>3</td><td>Read</td><td>Static Latency Stamp Unit [9]</td></tr>
<tr><td>8</td><td>2</td><td>De-allocate/TRIM</td><td>Static Latency Stamp Unit [8]</td></tr>
<tr><td>7</td><td>2</td><td>Write</td><td>Static Latency Stamp Unit [7]</td></tr>
<tr><td>6</td><td>2</td><td>Read</td><td>Static Latency Stamp Unit [6]</td></tr>
<tr><td>5</td><td>1</td><td>De-allocate/TRIM</td><td>Static Latency Stamp Unit [5]</td></tr>
<tr><td>4</td><td>1</td><td>Write</td><td>Static Latency Stamp Unit [4]</td></tr>
<tr><td>3</td><td>1</td><td>Read</td><td>Static Latency Stamp Unit [3]</td></tr>
<tr><td>2</td><td>0</td><td>De-allocate/TRIM</td><td>Static Latency Stamp Unit [2]</td></tr>
<tr><td>1</td><td>0</td><td>Write</td><td>Static Latency Stamp Unit [1]</td></tr>
<tr><td>0</td><td>0</td><td>Read</td><td>Static Latency Stamp Unit [0]</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-27</td><td>447:426</td><td>Reserved</td><td>22</td><td>Shall be cleared to zero.</td></tr>
<tr><td>LMDATA-28</td><td>449:448</td><td>Debug Log Trigger Enable</td><td>2</td><td>This controls what counters can cause a debug log event to be triggered. When set to 1b the first time the bucket/counter combination is incremented a debug log is triggered. When cleared to 0b a debug log will not be triggered when the bucket/counter combination is incremented.</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><th>Bit</th><th>Default Value</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>15:12</td><td>N/A</td><td>N/A</td><td>N/A</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>11</td><td>1b</td><td>3</td><td>De-allocate/TRIM</td><td>Active Log Enable [11]</td></tr>
<tr><td>10</td><td>1b</td><td>3</td><td>Write</td><td>Active Log Enable [10]</td></tr>
<tr><td>9</td><td>1b</td><td>3</td><td>Read</td><td>Active Log Enable [9]</td></tr>
<tr><td>8</td><td>1b</td><td>2</td><td>De-allocate/TRIM</td><td>Active Log Enable [8]</td></tr>
<tr><td>7</td><td>1b</td><td>2</td><td>Write</td><td>Active Log Enable [7]</td></tr>
<tr><td>6</td><td>1b</td><td>2</td><td>Read</td><td>Active Log Enable [6]</td></tr>
<tr><td>5</td><td>0b</td><td>1</td><td>De-allocate/TRIM</td><td>Active Log Enable [5]</td></tr>
<tr><td>4</td><td>0b</td><td>1</td><td>Write</td><td>Active Log Enable [4]</td></tr>
<tr><td>3</td><td>0b</td><td>1</td><td>Read</td><td>Active Log Enable [3]</td></tr>
<tr><td>2</td><td>0b</td><td>0</td><td>De-allocate/TRIM</td><td>Active Log Enable [2]</td></tr>
<tr><td>1</td><td>0b</td><td>0</td><td>Write</td><td>Active Log Enable [1]</td></tr>
<tr><td>0</td><td>0b</td><td>0</td><td>Read</td><td>Active Log Enable [0]</td></tr>
</table>
</td></tr>
<tr><td>LMDATA-29</td><td>451:450</td><td>Debug Log Measured Latency</td><td>2</td><td>When a debug log is triggered, this is the measured latency for the Latency Stamp that caused the debug log to trigger. A value of 0000h means this field is invalid. A value of</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>0001h represents 1ms. A value of 0002h represents 2 milliseconds. A value of FFFFh means the Debug Log Measured Latency saturated at the max.</td></tr>
<tr><td>LMDATA-30</td><td>459:452</td><td>Debug Log Latency Stamp</td><td>8</td><td>This is the Latency Stamp associated with the debug log.</td></tr>
<tr><td>LMDATA-31</td><td>461:460</td><td>Debug Log Pointer</td><td>2</td><td>If the debug log is valid this is a pointer to the debug log. This may be used as the telemetry identifier or some other type of pointer to the debug log.</td></tr>
<tr><td>LMDATA-32</td><td>463:462</td><td>Debug Counter Trigger Source</td><td>2</td><td>When the Debug Counter Trigger Source bit is set to 1b the debug log is valid, and this is the counter that triggered the debug log. When the Debug Counter Trigger Source bit is cleared to 0b this is not the counter that triggered the debug log. No more than 1 bit in this field shall be set. When the Debug Counter Trigger Source is 0b the Debug Log Latency, Debug Log Latency Stamp, Debug Log Pointer and Debug Counter Trigger Source is not valid.

If the debug log has been cleared this field shall be cleared to zero.
<table>
<tr><th>Bit</th><th>Bucket</th><th>Counter</th><th>Description</th></tr>
<tr><td>15:12</td><td>N/A</td><td>N/A</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>11</td><td>3</td><td>De-allocate/TRIM</td><td>Debug Counter Trigger Source 11</td></tr>
<tr><td>10</td><td>3</td><td>Write</td><td>Debug Counter Trigger Source 10</td></tr>
<tr><td>9</td><td>3</td><td>Read</td><td>Debug Counter</td></tr>
</table>
</td></tr>
</table>

Requirement ID	Byte Address	Field	# of Bytes	Field Description

			Trigger Source 9
8	2	De-allocate/TRIM	Debug Counter Trigger Source 8
7	2	Write	Debug Counter Trigger Source 7
6	2	Read	Debug Counter Trigger Source 6
5	1	De-allocate/TRIM	Debug Counter Trigger Source 5
4	1	Write	Debug Counter Trigger Source 4
3	1	Read	Debug Counter Trigger Source 3
2	0	De-allocate/TRIM	Debug Counter Trigger Source 2
1	0	Write	Debug Counter Trigger Source 1
0	0	Read	Debug Counter Trigger Source 0

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>LMDATA-33</td><td>464</td><td>Debug Log Stamp Units</td><td>1</td><td><table><tr><th>Bit</th><th>Description</th></tr><tr><td>7:1</td><td>Reserved. Shall be cleared to zero.</td></tr><tr><td>0</td><td>When set to 1b the Debug Latency Stamp was based on receiving the NVMe Timestamp and the offset from this. When cleared to 0b the Debug Latency Stamp was based on power on hours since the NVMe Timestamp was not received.</td></tr></table></td></tr>
<tr><td>LMDATA-34</td><td>493:465</td><td>Reserved</td><td>29</td><td>Shall be cleared to zero.</td></tr>
<tr><td>LMDATA-35</td><td>495:494</td><td>Log Page Version</td><td>2</td><td>This indicates the version of this log page. Shall be set to 0001h.</td></tr>
<tr><td>LMDATA-36</td><td>511:496</td><td>Log Page GUID</td><td>16</td><td>Shall be set to 85D45E58D4E643709C6C84D08CC07A92h.</td></tr>
</table>

4.8.10 Device Capabilities (Log Identifier C4h) Requirements

This log provides the host a consolidated report of critical device-specific support information. This vendor-specific log page, C4h shall be 4096 bytes with the following functional requirements and field format:

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>DCLP-1</td><td>1:0</td><td>PCI Express Ports</td><td>2</td><td>The number of physical PCI Express ports supported by the device.</td></tr>
<tr><td>DCLP-2</td><td>3:2</td><td>OOB Management Support</td><td>2</td><td>Bit field indicating the OOB Management interfaces supported by the device.<table><tr><th>Bit</th><th>Description</th></tr><tr><td>15</td><td>Shall be set to 1b if the device has been tested and found to comply with the OOB Management requirements of this specification.</td></tr><tr><td>14:3</td><td>Reserved. Shall be cleared to zero.</td></tr><tr><td>2</td><td>Shall be set to 1b if NVMe Basic Management Command is supported.</td></tr><tr><td>1</td><td>Shall be set to 1b if MCTP over PCIe VDM is supported.</td></tr></table></td></tr>
</table>

<table>
<tr><td></td><td></td><td></td><td></td><td><table><tr><td>0</td><td>Shall be set to 1b if MCTP over SMBus is supported.</td></tr></table></td></tr>
<tr><td>DCLP-3</td><td>5:4</td><td>Write Zeroes Command Support</td><td>2</td><td>Bit field indicating the Write Zeroes command requirements supported by the device.<table><tr><th>Bit</th><th>Description</th></tr><tr><td>15</td><td>Shall be set to 1b if the device has been tested and found to comply with the Write Zeroes command requirements of this specification.</td></tr><tr><td>14:5</td><td>Reserved. Shall be cleared to zero.</td></tr><tr><td>4</td><td>Shall be set to 1b if the requirements of NVMe-IO-6 (Write Zeroes DEAC bit) are met.</td></tr><tr><td>3</td><td>Shall be set to 1b if the requirements of NVMe-IO-5 are met.</td></tr><tr><td>2</td><td>Shall be set to 1b if setting the FUA bit is supported.</td></tr><tr><td>1</td><td>Shall be set to 1b if setting the DEAC bit is supported.</td></tr><tr><td>0</td><td>Shall be set to 1b if the Write Zeroes command is supported.</td></tr></table></td></tr>
<tr><td>DCLP-4</td><td>7:6</td><td>Sanitize Command Support</td><td>2</td><td>Bit field indicating the Sanitize command requirements supported by the device.<table><tr><th>Bit</th><th>Description</th></tr><tr><td>15</td><td>Shall be set to 1b if the device has been tested and found to comply with the Sanitize command requirements of this specification.</td></tr><tr><td>14:5</td><td>Reserved. Shall be cleared to zero.</td></tr><tr><td>4</td><td>Shall be set to 1b if Deallocate LBAs is supported.</td></tr><tr><td>3</td><td>Shall be set to 1b if Overwrite is supported.</td></tr><tr><td>2</td><td>Shall be set to 1b if Block Erase is supported.</td></tr><tr><td>1</td><td>Shall be set to 1b if Crypto-Erase is supported.</td></tr><tr><td>0</td><td>Shall be set to 1b if the Sanitize command is supported.</td></tr></table></td></tr>
</table>

<table>
<tr>
<td>DCLP-5</td>
<td>9:8</td>
<td>Dataset Management Command Support</td>
<td>2</td>
<td>Bit field indicating the Dataset Management command requirements supported by the device.
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>15</td><td>Shall be set to 1b if the device has been tested and found to comply with the Dataset Management command requirements of this specification.</td></tr>
<tr><td>14:2</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>1</td><td>Shall be set to 1b if Attribute – Deallocate (AD) is supported.</td></tr>
<tr><td>0</td><td>Shall be set to 1b if the Dataset Management command is supported.</td></tr>
</table>
</td>
</tr>
<tr>
<td>DCLP-6</td>
<td>11:10</td>
<td>Write Uncorrectable Command Support</td>
<td>2</td>
<td>Bit field indicating the Write Uncorrectable command requirements supported by the device.
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>15</td><td>Shall be set to 1b if the device has been tested and found to comply with the Write Uncorrectable command requirements of this specification.</td></tr>
<tr><td>14:4</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>3</td><td>Shall be set to 1b if the SMART / Health Information (Log Identifier 02h) requirements of NVMe-IO-14 are met.</td></tr>
<tr><td>2</td><td>Shall be set to 1b if the device supports setting uncorrectable on the maximum number of LBAs allowed by the NVMe specification.</td></tr>
<tr><td>1</td><td>Shall be set to 1b if the device supports setting uncorrectable on a single LBA.</td></tr>
<tr><td>0</td><td>Shall be set to 1b if the Write Uncorrectable command is supported.</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr><td>DCLP-7</td><td>13:12</td><td>Fused Operation Support</td><td>2</td><td>Bit field indicating the fused command pairs requirements supported by the device.<table><tr><th>Bit</th><th>Description</th></tr><tr><td>15</td><td>Shall be set to 1b if the device has been tested and found to comply with the fused command pair support requirements of this specification.</td></tr><tr><td>14:1</td><td>Reserved. Shall be cleared to zero.</td></tr><tr><td>0</td><td>Shall be set to 1b if the Compare and Write fused command pair is supported.</td></tr></table></td></tr>
<tr><td>DCLP-8</td><td>15:14</td><td>Minimum Valid DSSD Power State</td><td>2</td><td>Shall be set to the lowest numbered valid DSSD Power State. Setting a DSSD Power State less than this value results in Invalid Field in Command (see DSSDPSS-2 (DSSD Power State is less than the Minimum DSSD Power State)).</td></tr>
<tr><td>DCLP-9</td><td>143:16</td><td>DSSD Power State Descriptors</td><td>128</td><td>Byte 16 is reserved and shall be cleared to zero.

The device shall populate bytes 17-143 with a single-byte DSSD Power State Descriptor for each DSSD Power State (see Section 4.8.10.2 DSSD Power State Descriptor) starting with DSSD Power State 1.

The DSSD Power State Descriptors shall be populated in increasing order of power state number, i.e., byte 17 contains the descriptor for DSSD Power State 1, byte 18 contains the descriptor for DSSD Power State 2, etc. up to byte 143 which contains the descriptor for DSSD Power State 127.</td></tr>
<tr><td>DCLP-10</td><td>4077:144</td><td>Reserved</td><td>3934</td><td>Shall be cleared to zero.</td></tr>
<tr><td>DCLP-11</td><td>4079:4078</td><td>Log Page Version</td><td>2</td><td>This indicates the version of the mapping this log page uses. Shall be set to 0001h.</td></tr>
<tr><td>DCLP-12</td><td>4095:4080</td><td>Log Page GUID</td><td>16</td><td>Shall be set to B7053C914B58495D98C9E1D10D054297h.</td></tr>
</table>

4.8.10.1 DSSD Power State Requirements

DSSD power states provide an alternative interface to NVMe power management. DSSD and NVMe Power States are numbered in opposite directions – higher numbered DSSD Power States consume more

power than lower numbered DSSD Power States, whereas higher numbered NVMe Power States consume less power than lower numbered NVMe Power States.

Selecting an DSSD Power State via a Set Features command (see Section 4.12.13 DSSD Power State (Feature Identifier C7h) Set Feature) causes the device to run at the highest powered NVMe Power State whose Maximum Power (MP) in watts is less than or equal to the number of the DSSD Power State. For example, placing the device in DSSD Power State number 20 causes the device to enter the highest powered NVMe Power State whose Maximum Power (MP) is less than or equal to 20 watts.

In general, a device will have more DSSD Power States than NVMe Power States. Selecting an DSSD Power State whose number is not the Maximum Power (MP) in watts of an NVMe Power State causes the device to drop down to run at the highest powered NVMe Power State whose Maximum Power (MP) does not exceed the number of the DSSD Power State in watts.

For example, suppose that a device supports NVMe Power State 5 that consumes no more than 20 watts Maximum Power (MP) and NVMe Power State 6 that consumes no more than 16 watts Maximum Power (MP). Selecting DSSD Power State 18 (i.e., 18 watts Maximum Power (MP)) causes the device to enter NVMe Power State 6 and consume no more than 16 watts Maximum Power (MP). In this example the device reports 18 as its DSSD Power State and 6 as its NVMe Power State. Selecting NVMe Power State 6 causes the device to report 16 as its DSSD Power State and 6 as its NVMe Power State.

DSSD Power State 0 does not exist, as a device that is consuming 0 watts Maximum Power (MP) is powered off. DSSD Power States that represent a Maximum Power (MP) less than the lowest powered NVMe Power State are invalid. The lowest numbered valid DSSD Power State is indicated in DCLP-8.

4.8.10.2 DSSD Power State Descriptor

The DSSD Power State Descriptor for each DSSD Power State has the following format:

Requirement ID	Bits	Field	Field Description
DSSDPSD-1	7	Valid DSSD Power State	Shall be set to 1b if the number of this DSSD Power State is greater than or equal to the Minimum Valid OCP Power State (see DCLP-8 (Minimum Valid DSSD Power State)).
DSSDPSD-2	6:5	Reserved	Shall be cleared to zero.
DSSDPSD-3	4:0	NVMe Power State	Shall be set to the number of the highest powered NVMe Power State whose Maximum Power (MP) in watts is less than or equal to the number of this DSSD Power State.

4.8.11 Unsupported Requirements (Log Identifier C5h)

This log provides a host with a consolidated report of all the Requirement IDs that the device does not support. This vendor-specific log page, C5h shall be 4096 bytes with the following functional requirements and field format:

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>URLP-1</td><td>1:0</td><td>Unsupported Count</td><td>2</td><td>The number of Unsupported Requirement IDs.</td></tr>
<tr><td>URLP-2</td><td>15:2</td><td>Reserved</td><td>14</td><td>Shall be cleared to zero.</td></tr>
<tr><td>URLP-3</td><td>4063:16</td><td>Unsupported Requirements List</td><td>4048</td><td>This structure shall be populated with between 0 and 253 16-byte zero padded ASCII strings. Each string shall be the Requirement ID of a requirement that the device does not support. Entries shall be in alphabetical order. Unused entries shall be cleared to zero.
<table>
<tr><th>Byte Address</th><th>Description</th></tr>
<tr><td>31:16</td><td>First Unsupported Requirement ID entry.</td></tr>
<tr><td>47:32</td><td>Second Unsupported Requirement ID entry.</td></tr>
<tr><td>…</td><td>…</td></tr>
<tr><td>4063:4048</td><td>253rd Unsupported Requirement ID entry.</td></tr>
</table></td></tr>
<tr><td>URLP-4</td><td>4077:4064</td><td>Reserved</td><td>14</td><td>Shall be cleared to zero.</td></tr>
<tr><td>URLP-5</td><td>4079:4078</td><td>Log Page Version</td><td>2</td><td>This indicates the version of the mapping this log page uses. Shall be set to 0001h.</td></tr>
<tr><td>URLP-6</td><td>4095:4080</td><td>Log Page GUID</td><td>16</td><td>Shall be set to C7BB98B7D0324863BB2C23990E9C722Fh.</td></tr>
</table>

4.9 Firmware Update Requirements

This defines the requirements for firmware update in the device.

Requirement ID	Description
FWUP-1	A firmware activation history log shall be recorded (see Section 4.8.6 Firmware Activation History (Log Identifier C2h) Requirements).
FWUP-2	Devices shall not have any restrictions on the number of firmware downloads supported.
FWUP-3	The Firmware Commit command with the following Commit Action (CA) codes shall be supported: • 000b – Download only. • 001b – Download and activate upon reset. • 010b – Activate upon reset.

Requirement ID	Description
	• 011b – Activate immediately without reset.
FWUP-4	Firmware Image Download Command shall be supported.
FWUP-5	Obsolete.
FWUP-6	The device shall support a minimum of 2 read/write slots for firmware update and may support up to 7.
FWUP-7	For firmware commit action 011b (firmware activation without reset), the device shall complete the firmware activation process and be ready to accept host I/O and admin commands within 1 second from the receipt of the Firmware Commit command. The Maximum Time for Firmware Activation (MTFA) field shall not exceed Ah.
FWUP-8	When attempting to downgrade to incompatible firmware revision, the device shall return Firmware Activation Prohibited (13h) status to a Firmware Commit command. All lower security revision firmware images shall be considered incompatible.
FWUP-9	The firmware image in each valid firmware slot shall have multiple copies of that firmware image for reliability. A single corrupted firmware image shall not result in the device no longer functioning.
FWUP-10	Firmware activation shall not cause user data to be lost or destroyed.
FWUP-11	Firmware activation without reset shall preserve the current state of the device (e.g., Opal locking state, Set Features, Timestamp, Log Page contents, etc.).

4.10 De-Allocation Requirements

Requirement ID	Description
TRIM-1	The device shall support the Deallocate attribute of the Dataset Management command.
TRIM-2	For data that has been deallocated the NVMe specification version 1.4b requires it to be 0, 1, or unchanged when read. Data returned shall only be 0, 1 or unchanged on a sector-by-sector basis.
TRIM-3	If data has been de-allocated and not written to when an unsafe power down event happens, the data shall be 0, 1 or unchanged when read.
TRIM-4	De-allocated addresses shall provide the performance and reliability benefits of overprovisioned space.
TRIM-5	The device shall support Garbage Collection during periods of no I/O (Idle GC).
TRIM-6	Read latency shall not change more than 5% from baseline when the host is issuing De-Allocate/TRIM commands.

Requirement ID	Description
TRIM-7	Read latency shall not change more than 5% from baseline when the device is performing Idle GC.

4.11 Sector Size and Namespace Support

Requirement ID	Description
SECTOR-1	Obsolete. Replaced by SECTOR-4 (512-byte and 4096-byte logical block sizes).
SECTOR-2	Obsolete. Replaced by SECTOR-4 (512-byte and 4096-byte logical block sizes).
SECTOR-3	The device shall have one Namespace whose size is the maximum capacity as shipped from the factory.
SECTOR-4	The device shall support 512-byte and 4096-byte logical block sizes.

4.12 Set/Get Features Requirements

4.12.1 General Get Feature Requirements

Requirement ID	Description
GETF-1	For any Get Feature Identifier defined in this section, Selection (SEL) values 00b to 11b in Dword 10 shall be supported.
GETF-2	If the feature requested by Set Feature is not supported, then a status error code of 02h (Invalid Field in Command) shall be returned.

The device shall support the following additional vendor unique Set /Get Features Identifiers.

4.12.2 Error Injection (Feature Identifier C0h) Set Feature

Feature to inject one or more error conditions to be reported by the device. If multiple Set Features commands for this feature are sent by the host, then only information from the most recent successful command is retained (i.e., subsequent commands replace information provided by previous commands).

Requirement ID	Dword	Field	Bits	Field Description
SERRI-1	0	Command Identifier (CID)	31:16	Shall be set as defined in NVMe Specification version 1.4b.
SERRI-2	0	PRP or SGL for Data	15:14	Shall be cleared to 00b.

Requirement ID	Dword	Field	Bits	Field Description
		Transfer (PSDT)		
SERRI-3	0	Reserved	13:10	Shall be cleared to zero.
SERRI-4	0	Fused Operation (FUSE)	9:8	Shall be cleared to 00b.
SERRI-5	0	Opcode (OPC)	7:0	Shall be set to 09h.
SERRI-6	1	Namespace Identifier (NSID)	31:0	Shall be set to FFFFFFFFh.
SERRI-7	2:3	Reserved	31:0	Shall be cleared to zero.
SERRI-8	4:5	Metadata Pointer (MPTR)	31:0	Shall be cleared to zero.
SERRI-9	6:9	Data Pointer (DPTR)	31:0	Shall point to a physically contiguous 4096-byte address range containing 0 to 127 Error Injection Data Structure Entries.
SERRI-10	10	Save (SV)	31	The device shall not support setting this bit to 1b. If the controller receives this Set Features command with the bit set to 1b, then the device shall abort the command with a status of Feature Identifier Not Saveable.
SERRI-11	10	Reserved	30:8	Shall be cleared to zero.
SERRI-12	10	Feature Identifier (FID)	7:0	Shall be set to C0h.
SERRI-13	11	Reserved	31:7	Shall be cleared to zero.
SERRI-14	11	Number of Error Injections	6:0	This field shall specify the number of valid Error Injection Data Entries described in the address range pointed to by the Data Pointer (DPTR) field.
SERRI-15	12:13	Reserved	31:0	Shall be cleared to zero.
SERRI-16	14	UUID Index	31:0	Shall be set per UUID-3.
SERRI-17	15	Reserved	31:0	Shall be cleared to zero.

Requirement ID	Description
ERRI-1	The maximum number of entries in the Number of Error Injections field shall be 127.
ERRI-2	A value of 0000000b in the Number of Error Injections field shall clear any outstanding error injection events.
ERRI-3	The error injections shall not overlap and may be listed in any order (e.g., ordering by error injection type is not required).
ERRI-4	The host shall clear any unused entries in the Error Injection data structure to zero and the device shall ignore all zeroed entries. The device shall check at least the first four bytes of each Error Injection Entry data structure to determine if it is zeroed.
ERRI-5	The device shall abort the Error Injection Set Feature command if the request contains an error injection type that is not supported or the Single Instance value for the given Error Injection Type is not valid.
ERRI-6	Once the trigger conditions specified in an Error Injection Entry are met, the device shall inject the defined error event such that the host can detect the error through either an AEN being sent, the CFS bit being set, or command being aborted.

4.12.2.1 Error Injection Entry Data Structure

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>ERRIE-1</td><td>0</td><td>Error Entry Flags</td><td>1</td><td>Error Entry Flags definition:
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>7:2</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>1</td><td>Single Instance: If cleared to 0b, indicates error injection is enabled until disabled. If set to 1b, indicates a single instance error injection where a single error shall be injected. After a single instance error has been created, the error injection is considered disabled.</td></tr>
<tr><td>0</td><td>Error Injection Enable: If cleared to 0b, indicates error injection is disabled. If set to 1b, indicates error injection is enabled.</td></tr>
</table></td></tr>
<tr><td>ERRIE-2</td><td>1</td><td>Reserved</td><td>1</td><td>Shall be cleared to zero.</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>ERRIE-3</td><td>3:2</td><td>Error Injection Type</td><td>2</td><td>Error Injection type definition:
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>0000h</td><td>Reserved.</td></tr>
<tr><td>0001h</td><td>Device Panic – CPU/Controller Hang</td></tr>
<tr><td>0002h</td><td>Device Panic – NAND Hang</td></tr>
<tr><td>0003h</td><td>Device Panic – PLP Defect</td></tr>
<tr><td>0004h</td><td>Device Panic – Logical Firmware Error</td></tr>
<tr><td>0005h</td><td>Device Panic – DRAM Corruption Critical Path</td></tr>
<tr><td>0006h</td><td>Device Panic – DRAM Corruption Non-Critical Path</td></tr>
<tr><td>0007h</td><td>Device Panic – NAND Corruption</td></tr>
<tr><td>0008h</td><td>Device Panic – SRAM Corruption</td></tr>
<tr><td>0009h</td><td>Device Panic – HW Malfunction</td></tr>
<tr><td>000Ah</td><td>Device Panic – No More NAND Spares Available</td></tr>
<tr><td>000Bh - FFFFh</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table></td></tr>
<tr><td>ERRIE-4</td><td>31:4</td><td>Error Injection Type Specific Definition</td><td>28</td><td>Error Injection Type specific definition.</td></tr>
</table>

4.12.2.2 Device Panic Error Injection Type

The device shall inject a device panic that the host can detect through either an AEN or the CFS bit being set. For the Device Panic type, a Single Instance value of 0 is not valid. Host shall perform the Panic Reset and Device Recovery actions specified in Error Recovery (Log Identifier C1h).

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>ERRIEDP-1</td><td>0</td><td>Error Entry Flags</td><td>1</td><td>Device Panic Error Entry Flags:
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>7:2</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>1</td><td>Shall be set to 1b.</td></tr>
<tr><td>0</td><td>Shall be set to 1b.</td></tr>
</table></td></tr>
<tr><td>ERRIEDP-2</td><td>1</td><td>Reserved</td><td>1</td><td>Shall be cleared to zero.</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>ERRIEDP-3</td><td>3:2</td><td>Error Injection Type</td><td>2</td><td>Shall be set to the range of 0001h to 000Ah.</td></tr>
<tr><td>ERRIEDP-4</td><td>31:4</td><td>Error Injection Type Specific Definition</td><td>28</td><td>Device Panic Error Injection information:
<table>
<tr><th>Byte Address</th><th>Field Description</th></tr>
<tr><td>31:6</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>5:4</td><td>Number of Reads to Trigger Device Panic (NRTDP): Indicates the number of Read commands the device shall process and complete before triggering a device panic.</td></tr>
<tr><td></td><td></td></tr>
</table></td></tr>
</table>

4.12.3 Error Injection (Feature Identifier C0h) Get Feature

This Get Feature returns the set of error injections that are enabled on the device. The attributes specified in Section 4.12.2.2 - Device Panic Error Injection Type are returned in Dword 0 of the completion queue entry and the Error Inject data structure specified in Section 4.12.2.1 - Error Injection Entry Data Structure is returned for each error injection in the data buffer for that command. If there are no currently enabled error injections, the data buffer returned shall contain all zeros. The device shall clear to zero all unused entries in the Error Injection data structure.

4.12.3.1 Error Injection – Get Features Completion Queue Entry Dword 0

Requirement ID	Field	Bits	Field Description
GERRI-1	Reserved	31:7	Shall be cleared to zero.
GERRI-2	Number of Error Injections (NUM)	6:0	This field indicates the number of outstanding enabled error injections returned in the command data buffer (see Section 4.12.2.1 Error Injection Entry Data Structure for the format of the entries).

4.12.4 Clear Firmware Update History (Feature Identifier C1h) Set Feature

Requirement ID	Dword	Field	Bits	Field Description
CFUH-1	0	Command Identifier (CID)	31:16	Shall be set as defined in NVMe Specification version 1.4b.
CFUH-2	0	PRP or SGL for Data Transfer (PSDT)	15:14	Shall be cleared to 00b.
CFUH-3	0	Reserved	13:10	Shall be cleared to zero.
CFUH-4	0	Fused Operation (FUSE)	9:8	Shall be cleared to 00b.
CFUH-5	0	Opcode (OPC)	7:0	Shall be set to 09h.
CFUH-6	1	Namespace Identifier (NSID)	31:0	Shall be cleared to zero.
CFUH-7	2:3	Reserved	31:0	Shall be cleared to zero.
CFUH-8	4:5	Metadata Pointer (MPTR)	31:0	Shall be cleared to zero.
CFUH-9	6:9	Data Pointer (DPTR)	31:0	Shall be cleared to zero.
CFUH-10	10	Save (SV)	31	The device shall not support setting this bit to 1b. If the controller receives this Set Features command with the bit set to 1b, then the device shall abort the command with a status of Feature Identifier Not Saveable.
CFUH-11	10	Reserved	30:8	Shall be cleared to zero.
CFUH-12	10	Feature Identifier (FID)	7:0	Shall be set to C1h.
CFUH-13	11	Clear Firmware Update History Log	31	Set to 1b to clear the Firmware Activation History (Log Identifier C2h). The NVMe CLI plug in command “clear-fw-activate-history” can also perform this operation.
CFUH-14	11	Reserved	30:0	Shall be cleared to zero.

Requirement ID	Dword	Field	Bits	Field Description
CFUH-15	12:13	Reserved	31:0	Shall be cleared to zero.
CFUH-16	14	UUID Index	31:0	Shall be set per UUID-3.
CFUH-17	15	Reserved	31:0	Shall be cleared to zero.

4.12.5 EOL/PLP Failure Mode (Feature Identifier C2h) Set Feature

This Set Feature defines the mode to which the device shall transition at End of Life (EOL) or on failure of the Power Loss Protection (PLP) circuitry.

Requirement ID	Description
ROWTM-1	The device shall default from the factory to Read Only Mode (ROM) (01b).

Requirement ID	Dword	Field	Bits	Field Description
SROWTM-1	0	Command Identifier (CID)	31:16	Shall be set as defined in NVMe Specification version 1.4b.
SROWTM-2	0	PRP or SGL for Data Transfer (PSDT)	15:14	Shall be cleared to 00b.
SROWTM-3	0	Reserved	13:10	Shall be cleared to zero.
SROWTM-4	0	Fused Operation (FUSE)	9:8	Shall be cleared to 00b.
SROWTM-5	0	Opcode (OPC)	7:0	Shall be set to 09h.
SROWTM-6	1	Namespace Identifier (NSID)	31:0	Shall be cleared to zero.
SROWTM-7	2:3	Reserved	31:0	Shall be cleared to zero.
SROWTM-8	4:5	Metadata Pointer (MPTR)	31:0	Shall be cleared to zero.
SROWTM-9	6:9	Data Pointer (DPTR)	31:0	Shall be cleared to zero.
SROWTM-10	10	Save (SV)	31	This bit specifies that the controller shall save the End-of-Life Behavior state so that the state

<table>
<tr><th>Requirement ID</th><th>Dword</th><th>Field</th><th>Bits</th><th>Field Description</th></tr>
<tr><td></td><td></td><td></td><td></td><td>persists through all power states and resets. The device shall support setting this bit to 1b.</td></tr>
<tr><td>SROWTM-11</td><td>10</td><td>Reserved</td><td>30:8</td><td>Shall be cleared to zero.</td></tr>
<tr><td>SROWTM-12</td><td>10</td><td>Feature Identifier (FID)</td><td>7:0</td><td>Shall be set to C2h.</td></tr>
<tr><td>SROWTM-13</td><td>11</td><td>End of Life Behavior</td><td>31:30</td><td>Field to indicate device write behavior at End of Life (EOL) or in the event of loss of PLP functionality. See EOL-5 (available spares) for definition on EOL.
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>00b</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>01b</td><td>The device shall transition to Read Only Mode (ROM) in the event of PLP failure or at EOL.</td></tr>
<tr><td>10b</td><td>The device shall transition to Write Through Mode (WTM) in the event of PLP failure and transition to Read Only Mode (ROM) at EOL.</td></tr>
<tr><td>11b</td><td>The device shall continue to operate as normal in the event of PLP failure and transition to Read Only Mode (ROM) at EOL.</td></tr>
</table></td></tr>
<tr><td>SROWTM-14</td><td>11</td><td>Reserved</td><td>29:0</td><td>Shall be cleared to zero.</td></tr>
<tr><td>SROWTM-15</td><td>12:15</td><td>Reserved</td><td>31:0</td><td>Shall be cleared to zero.</td></tr>
<tr><td>SROWTM-16</td><td>14</td><td>UUID Index</td><td>31:0</td><td>Shall be set per UUID-3.</td></tr>
<tr><td>SROWTM-17</td><td>15</td><td>Reserved</td><td>31:0</td><td>Shall be cleared to zero.</td></tr>
</table>

4.12.6 EOL/PLP Failure Mode (Feature Identifier C2h) Get Feature

Dword 0 of command completion queue entry.

Requirement ID	Field	Bits	Field Description
GROWTM-1	Reserved	31:3	Shall be cleared to zero.

<table>
<tr><th>Requirement ID</th><th>Field</th><th>Bits</th><th>Field Description</th></tr>
<tr><td>GROWTM -2</td><td>End of Life Behavior</td><td>2:0</td><td>Field to indicate what the device write behavior is configured for at End of Life (EOL) or in the event of loss of PLP functionality.
The tables below define the required return values for each Selection (SEL) state. All other bit values are reserved.

Current state (Selection (SEL) cleared to 00b):
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>001b</td><td>The device will transition to Read Only Mode (ROM) in the event of PLP failure or at EOL.</td></tr>
<tr><td>010b</td><td>The device will transition to Write Through Mode (WTM) in the event of PLP failure and transition to Read Only Mode (ROM) at EOL.</td></tr>
<tr><td>011b</td><td>The device will continue to operate as normal in the event of PLP failure and transition to Read Only Mode (ROM) at EOL.</td></tr>
</table>
Default state (Selection (SEL) set to 01b):
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>001b</td><td>Read Only Mode (ROM) is the factory default.</td></tr>
<tr><td>010b</td><td>The Write Through Mode (WTM) is the factory default.</td></tr>
<tr><td>011b</td><td>Normal operation is the factory default for PLP failure. Read Only Mode (ROM) is the factory default at EOL.</td></tr>
</table>
Saved state (Selection (SEL) set to 10b):
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>001b</td><td>The saved state is Read Only Mode (ROM).</td></tr>
<tr><td>010b</td><td>The saved state is Write Through Mode (WTM).</td></tr>
<tr><td>011b</td><td>The saved state is to operate as normal in the event of PLP failure and Read Only Mode (ROM) at EOL.</td></tr>
</table>
Capabilities (Selection (SEL) set to 11b):
<table>
<tr><th>Value</th><th>Field Description</th></tr>
</table></td></tr>
</table>

Requirement ID	Field	Bits	Field Description	
			101b	This feature is saveable, changeable, and not namespace specific.

4.12.7 Clear PCIe Correctable Error Counters (Feature Identifier C3h) Set Feature

Requirement ID	Dword	Field	Bits	Field Description
CPCIE-1	0	Command Identifier (CID)	31:16	Shall be set as defined in NVMe Specification version 1.4b.
CPCIE -2	0	PRP or SGL for Data Transfer (PSDT)	15:14	Shall be cleared to 00b.
CPCIE -3	0	Reserved	13:10	Shall be cleared to zero.
CPCIE -4	0	Fused Operation (FUSE)	9:8	Shall be cleared to 00b.
CPCIE-5	0	Opcode (OPC)	7:0	Shall be set to 09h.
CPCIE-6	1	Namespace Identifier (NSID)	31:0	Shall be cleared to zero.
CPCIE-7	2:3	Reserved	31:0	Shall be cleared to zero.
CPCIE-8	4:5	Metadata Pointer (MPTR)	31:0	Shall be cleared to zero.
CPCIE-9	6:9	Data Pointer (DPTR)	31:0	Shall be cleared to zero.
CPCIE-10	10	Save (SV)	31	The device shall not support setting this bit to 1b. If the controller receives this Set Features command with the bit set to 1b, then the device shall abort the command with a status of Feature Identifier Not Saveable.
CPCIE-11	10	Reserved	30:8	Shall be cleared to zero.
CPCIE-12	10	Feature Identifier (FID)	7:0	Shall be set to C3h.

Requirement ID	Dword	Field	Bits	Field Description
CPCIE-13	11	Clear PCIe Error Counters	31	Set to 1b to clear all PCIe correctable error counters in the SMART / Health Information Extended (Log Identifier C0h). The NVMe CLI plug-in command “clear-pcie-correctable-errors” can also perform this operation.
CPCIE-14	11	Reserved	30:0	Shall be cleared to zero.
CPCIE-15	12:13	Reserved	31:0	Shall be cleared to zero.
CPCIE-16	14	UUID Index	31:0	Shall be set per UUID-3.
CPCIE-17	15	Reserved	31:0	Shall be cleared to zero.

4.12.8 Enable IEEE1667 Silo (Feature Identifier C4h) Set Feature

This Set Feature shall return an error if the OPAL Security state is not Manufactured_Inactive.

Requirement ID	Dword	Field	Bits	Field Description
S1667-1	0	Command Identifier (CID)	31:16	Shall be set as defined in NVMe Specification version 1.4b.
S1667-2	0	PRP or SGL for Data Transfer (PSDT)	15:14	Shall be cleared to 00b.
S1667-3	0	Reserved	13:10	Shall be cleared to zero.
S1667-4	0	Fused Operation (FUSE)	9:8	Shall be cleared to 00b.
S1667-5	0	Opcode (OPC)	7:0	Shall be set to 09h.
S1667-6	1	Namespace Identifier (NSID)	31:0	Shall be cleared to zero.
S1667-7	2:3	Reserved	31:0	Shall be cleared to zero.
S1667-8	4:5	Metadata Pointer (MPTR)	31:0	Shall be cleared to zero.
S1667-9	6:9	Data Pointer (DPTR)	31:0	Shall be cleared to zero.

Requirement ID	Dword	Field	Bits	Field Description
S1667-10	10	Save (SV)	31	This bit specifies that the controller shall save the IEEE1667 Silo Enable/Disable state so that the state persists through all power states and resets. The device shall support setting this bit to 1b.
S1667-11	10	Reserved	30:8	Shall be cleared to zero.
S1667-12	10	Feature Identifier (FID)	7:0	Shall be set to C4h.
S1667-13	11	Enable IEEE1667 Silo	31	If set to 0b, the IEEE 1667 silo shall be disabled no later than the next power cycle. If set to 1b, the IEEE 1667 silo shall be enabled no later than the next power cycle.
S1667-14	11	Reserved	30:0	Shall be cleared to zero.
S1667-15	12:13	Reserved	31:0	Shall be cleared to zero.
S1667-16	14	UUID Index	31:0	Shall be set per UUID-3.
S1667-17	15	Reserved	31:0	Shall be cleared to zero.

4.12.9 Enable IEEE1667 Silo (Feature Identifier C4h) Get Feature

Dword 0 of command completion queue entry.

<table>
<tr><th>Requirement ID</th><th>Field</th><th>Bits</th><th>Field description</th></tr>
<tr><td>G1667-1</td><td>Reserved</td><td>31:3</td><td>Shall be cleared to zero.</td></tr>
<tr><td>G1667-2</td><td>IEEE1667 Silo Enabled</td><td>2:0</td><td>The tables below define the required return values for each Selection (SEL) state. All other values are illegal.

Current state (Selection (SEL) cleared to 00b):
<table><tr><th>Value</th><th>Field Description</th></tr><tr><td>000b</td><td>The IEEE1667 silo is currently disabled.</td></tr><tr><td>001b</td><td>The IEEE1667 silo is currently enabled.</td></tr></table>
Default state (Selection (SEL) set to 01b):
<table><tr><th>Value</th><th>Field Description</th></tr><tr><td>000b</td><td>The IEEE1667 silo factory default is disabled.</td></tr><tr><td>001b</td><td>The IEEE1667 silo factory default is enabled.</td></tr></table>
Saved state (Selection (SEL) set to 10b):</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Field</th><th>Bits</th><th>Field description</th></tr>
<tr><td></td><td></td><td></td><td>
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>000b</td><td>The IEEE1667 silo saved state is disabled.</td></tr>
<tr><td>001b</td><td>The IEEE1667 silo saved state is enabled.</td></tr>
</table>
Capabilities (Selection (SEL) set to 11b):
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>101b</td><td>This feature is saveable, changeable, and not namespace specific.</td></tr>
</table>
</td></tr>
</table>

4.12.10 Latency Monitor (Feature Identifier C5h) Set Feature

This configures the Latency Monitor Feature.

Requirement ID	Dword	Field	Bits	Field description
LMSF-1	0	Command Identifier (CID)	31:16	Shall be set as defined in NVMe Specification version 1.4b.
LMSF-2	0	PRP or SGL for Data Transfer (PSDT)	15:14	Shall be cleared to 00b.
LMSF-3	0	Reserved	13:10	Shall be cleared to zero.
LMSF-4	0	Fused Operation (FUSE)	9:8	Shall be cleared to 00b.
LMSF-5	0	Opcode (OPC)	7:0	Shall be set to 09h.
LMSF-6	1	Namespace Identifier (NSID)	31:0	Shall be cleared to zero.
LMSF-7	3:2	Reserved	31:0	Shall be cleared to zero.
LMSF-8	5:4	Metadata Pointer (MPTR)	31:0	Shall be cleared to zero.
LMSF-9	9:6	Data Pointer (DPTR)	31:0	Shall point to a physically contiguous 4096-byte address range (see Section 4.12.10.1 Latency Monitoring Data Structure Entry).

Requirement ID	Dword	Field	Bits	Field description
LMSF-10	10	Save (SV)	31	This bit specifies that the controller shall save the data pointed to by the Data Pointer in LMSF-9 (Data Pointer (DPTR)) so that the data persists through all power states and resets. The device shall support setting this bit to 1b.
LMSF-11	10	Reserved	30:8	Shall be cleared to zero.
LMSF-12	10	Feature Identifier (FID)	7:0	Shall be set to C5h.
LMSF-13	11	Reserved	31:0	Shall be cleared to zero.
LMSF-14	12:13	Reserved	31:0	Shall be cleared to zero.
LMSF-15	14	UUID Index	31:0	Shall be set per UUID-3.
LMSF-16	15	Reserved	31:0	Shall be cleared to zero.

4.12.10.1 Latency Monitoring Data Structure Entry

This data structure is 4096 bytes with the following functional requirements and field format:

Requirement ID	Byte Address	Field	# of Bytes	Field Description
LMDS-1	1:0	Active Bucket Timer Threshold	2	This is the value that loads the Active Bucket Timer Threshold.
LMDS-2	2	Active Threshold A	1	This is the value that loads into the Active Threshold A
LMDS-3	3	Active Threshold B	1	This is the value that loads into the Active Threshold B
LMDS-4	4	Active Threshold C	1	This is the value that loads into the Active Threshold C
LMDS-5	5	Active Threshold D	1	This is the value that loads into the Active Threshold D
LMDS-6	7:6	Active Latency Config.	2	This is the value that loads into the Active Latency Configuration.
LMDS-7	8	Active Latency Minimum Window	1	This is the value that loads into the Active Latency Minimum Window.

<table>
<tr><th>Requirement ID</th><th>Byte Address</th><th>Field</th><th># of Bytes</th><th>Field Description</th></tr>
<tr><td>LMDS-8</td><td>10:9</td><td>Debug Log Trigger Enable</td><td>2</td><td>This is the value that loads into the Debug Log Trigger Enable.</td></tr>
<tr><td>LMDS-9</td><td>11</td><td>Discard Debug Log</td><td>1</td><td>
<table>
<tr><th>Value</th><th>Description</th></tr>
<tr><td>00h</td><td>When cleared to 00h the debug log, if it exists, will not be cleared.</td></tr>
<tr><td>01h</td><td>When set to 01h the debug log will be discarded so another log can be triggered.

In addition to this all the fields in the Set Features Data structure are valid which will cause the Latency Monitor feature to be reset and loaded with the values from the Set Features command.

This shall discard the Debug Log regardless of whether the Latency Monitoring Feature is enabled or disabled.</td></tr>
<tr><td>02h</td><td>When set to 02h the debug log will be discarded so another log can be triggered.

In addition to this none of the other fields of the Set Features Data structure are valid. Thus, only the debug log is discarded, and the Latency Monitor feature is not reset or loaded with any new values from the Set Features command.

This shall discard the Debug Log regardless of whether the Latency Monitoring Feature is enabled or disabled.</td></tr>
<tr><td>03h -FFh</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table>
</td></tr>
</table>

Requirement ID	Byte Address	Field	# of Bytes	Field Description
LMDS-10	12	Latency Monitor Feature Enable	1	When set to 01h the Latency Monitor Feature is enabled. When cleared to 00h the Latency Monitor Feature is disabled. All other values are reserved.
LMDS-11	4095:13	Reserved	4083	Shall be cleared to zero.

4.12.11 PLP Health Check Interval (Feature Identifier C6h) Set Feature

This Set Feature defines the test interval that the device shall use to test that it's Power Loss Protection (PLP) circuitry is healthy.

Requirement ID	Dword	Field	Bits	Field Description
SPLPI-1	0	Command Identifier (CID)	31:16	Shall be set as defined in NVMe Specification version 1.4b.
SPLPI-2	0	PRP or SGL for Data Transfer (PSDT)	15:14	Shall be cleared to 00b.
SPLPI-3	0	Reserved	13:10	Shall be cleared to zero.
SPLPI-4	0	Fused Operation (FUSE)	9:8	Shall be cleared to 00b.
SPLPI-5	0	Opcode (OPC)	7:0	Shall be set to 09h.
SPLPI-6	1	Namespace Identifier (NSID)	31:0	Shall be cleared to zero.
SPLPI-7	2:3	Reserved	31:0	Shall be cleared to zero.
SPLPI-8	4:5	Metadata Pointer (MPTR)	31:0	Shall be cleared to zero.
SPLPI-9	6:9	Data Pointer (DPTR)	31:0	Shall be cleared to zero.
SPLPI-10	10	Save (SV)	31	This bit specifies that the controller shall save the PLP Health Check Interval so that the time interval persists through all power states and

Requirement ID	Dword	Field	Bits	Field Description
				resets. The device shall support setting this bit to 1b.
SPLPI-11	10	Reserved	30:8	Shall be cleared to zero.
SPLPI-12	10	Feature Identifier (FID)	7:0	Shall be set to C6h.
SPLPI-13	11	PLP Health Check Interval	31:16	The time interval between PLP health checks in minutes. If cleared to 0000h, then the health check is disabled.
SPLPI-14	11	Reserved	15:0	Shall be cleared to zero.
SPLPI-15	12:13	Reserved	31:0	Shall be cleared to zero.
SPLPI-16	14	UUID Index	31:0	Shall be set per UUID-3.
SPLPI-17	15	Reserved	31:0	Shall be cleared to zero.

4.12.12 PLP Health Check Interval (Feature Identifier C6h) Get Feature

Dword 0 of command completion queue entry.

Requirement ID	Field	Bits	Field description
GPLPI-1	Reserved	31:16	Shall be cleared to zero.
GPLPI-2	PLP Health Check Interval	15:0	The tables below define the required return values for each Selection (SEL) state. All other values are illegal. Current state (Selection (SEL) cleared to 00b): **Value** / **Field Description** 0000h: The PLP Health Check is currently disabled. xxxxh: The PLP Health Check Interval is currently xxxxh minutes. Default state (Selection (SEL) set to 01b): **Value** / **Field Description** 0000h: The PLP Health Check Interval factory default is disabled. 000Fh: The PLP Health Check Interval factory default is 000Fh minutes. Saved state (Selection (SEL) set to 10b):

<table>
<tr><th>Requirement ID</th><th>Field</th><th>Bits</th><th>Field description</th></tr>
<tr><td></td><td></td><td></td><td>
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>0000h</td><td>The PLP Health Check Interval saved state is disabled.</td></tr>
<tr><td>xxxxh</td><td>The PLP Health Check Interval saved state is xxxxh minutes.</td></tr>
</table>
Capabilities (Selection (SEL) set to 11b):
<table>
<tr><th>Value</th><th>Field Description</th></tr>
<tr><td>0005h</td><td>This feature is saveable, changeable, and not namespace specific.</td></tr>
</table>
</td></tr>
</table>

4.12.13 DSSD Power State (Feature Identifier C7h) Set Feature

Requirement ID	Description
DSSDPSS-1	If the host selects an DSSD Power State via a DSSD Power State (Feature Identifier C7h) Set Feature, the device shall accept the command and run at the highest powered NVMe Power State whose Maximum Power (MP) is less than or equal to the number of the DSSD Power State in watts (see DCLP-9 (DSSD Power State Descriptors) and DSSDPSD-3 (NVMe Power State)).
DSSDPSS-2	If the host selects a DSSD Power State via DSSD Power State (Feature Identifier C7h) Set Feature and the number of that DSSD Power State is less than the Minimum DSSD Power State (see DCLP-8 (Minimum Valid DSSD Power State)), the device shall abort the Set Feature command with Invalid Field in Command status.

This Set Feature causes the device to move the given DSSD Power State:

Requirement ID	Dword	Field	Bits	Field Description
SDSSDPS-1	0	Command Identifier (CID)	31:16	Shall be set as defined in NVMe Specification version 1.4b.
SDSSDPS-2	0	PRP or SGL for Data Transfer (PSDT)	15:14	Shall be cleared to 00b.
SDSSDPS-3	0	Reserved	13:10	Shall be cleared to zero.

Requirement ID	Dword	Field	Bits	Field Description
SDSSDPS-4	0	Fused Operation (FUSE)	9:8	Shall be cleared to 00b.
SDSSDPS-5	0	Opcode (OPC)	7:0	Shall be set to 09h.
SDSSDPS-6	1	Namespace Identifier (NSID)	31:0	Shall be cleared to zero.
SDSSDPS-7	2:3	Reserved	31:0	Shall be cleared to zero.
SDSSDPS-8	4:5	Metadata Pointer (MPTR)	31:0	Shall be cleared to zero.
SDSSDPS-9	6:9	Data Pointer (DPTR)	31:0	Shall be cleared to zero.
SDSSDPS-10	10	Save (SV)	31	This bit specifies that the controller shall persist the DSSD Power State through all power cycles and resets. The device shall support setting this bit to 1b.
SDSSDPS-11	10	Reserved	30:8	Shall be cleared to zero.
SDSSDPS-12	10	Feature Identifier (FID)	7:0	Shall be set to C7h.
SDSSDPS-13	11	Reserved	31:7	Shall be cleared to zero.
SDSSDPS-14	11	DSSD Power State	6:0	DSSD Power State to set in watts.
SDSSDPS-15	12:13	Reserved	31:0	Shall be cleared to zero.
SDSSDPS-16	14	UUID Index	31:0	Shall be set per UUID-3.
SDSSDPS-17	15	Reserved	31:0	Shall be cleared to zero.

4.12.14 DSSD Power State (Feature Identifier C7h) Get Feature

Requirement ID	Description
DSSDPSG-1	If an DSSD Power State (Feature Identifier C7h) Get Feature command is executed and the current NVMe Power State was selected because of a DSSD Power State Set Feature command, the device shall report the DSSD Power State selected by that command (see Section 4.8.10.1 DSSD Power State Requirements for an example).

Requirement ID	Description
DSSDPSG-2	If a DSSD Power State (Feature Identifier C7h) Get Feature command is executed and the current NVMe Power State was not selected because of a DSSD Power State Set Feature command, the device shall report the lowest power DSSD Power State whose number is greater than or equal to the Maximum Power in watts of the current NVMe Power State (see Section 4.8.10.1 DSSD Power State Requirements for an example).

Dword 0 of command completion queue entry.

Requirement ID	Field	Bits	Field description
GDSSDPS-1	Reserved	31:7	Shall be cleared to zero.
GDSSDPS-2	DSSD Power State	6:0	The tables below define the required return values for each Selection (SEL) state. All other values are illegal. Current state (Selection (SEL) cleared to 00b): **Value** / **Field Description** xx / The DSSD Power State is currently xx watts as specified by DSSDPSG-1 and DSSDPSG-2. Default state (Selection (SEL) set to 01b): **Value** / **Field Description** xx / The DSSD Power State factory default is xx watts. Saved state (Selection (SEL) set to 10b): **Value** / **Field Description** xx / The DSSD Power State saved state is xx watts. Capabilities (Selection (SEL) set to 11b): **Value** / **Field Description** 0005h / This feature is saveable, changeable, and not namespace specific.

5 PCIe Requirements

The following are PCIe requirements.

Requirement ID	Description
PCI-1	The device shall support a PCIe Maximum Payload Size (MPS) of 256 bytes or larger.
PCI-2	The device Controller shall support modification of PCIe TLP completion timeout range as defined by the PCIe Base Spec.
PCI-3	The vendor shall disclose the vendor-specific timeout range definition in the "Completion Timeout Ranges Supported" field in the PCI Express Base Specification.
PCI-4	Disabling of PCIe Completion Timeout shall also be supported by the device Controller.
PCI-5	All three PCIe Conventional Resets (Cold, Warm, Hot) shall be supported.
PCI-6	PCIe Function Level Reset shall be supported.
PCI-7	Obsolete. See PCI-5 (PCIe Conventional Resets).

5.1 Boot Requirements

Requirement ID	Description
BOOT-1	The device shall support the UEFI 2.0 and later in-box NVMe driver.
BOOT-2	An option ROM shall not be included.

5.2 PCIe Error Logging

The following table indicates where the PCIe physical layer error counters shall be logged. This is in addition to the aggregated PCIe error counters defined in SMART / Health Information Extended (Log Identifier C0h).

Requirement ID	Event	Logging Mechanism
PCIERR-1	Unsupported Request Error Status (URES)	Uncorrectable Error Status Register Offset 04h in PCIe Base Specification
PCIERR-2	ECRC Error Status (ECRCES)	
PCIERR-3	Malformed TLP Status (MTS)	
PCIERR-4	Receiver Overflow Status (ROS)	
PCIERR-5	Unexpected Completion Status (UCS)	
PCIERR-6	Completer Abort Status (CAS)	
PCIERR-7	Completion Timeout Status (CTS)	
PCIERR-8	Flow Control Protocol Error Status (FCPES)	
PCIERR-9	Poisoned TLP Status (PTS)	

Requirement ID	Event	Logging Mechanism
PCIERR-10	Data Link Protocol Error Status (DLPES)	
PCIERR-11	Advisory Non-Fatal Error Status (ANFES)	Correctable Error Status Register Offset 10h in PCIe Base Specification
PCIERR-12	Replay Timer Timeout Status (RTS)	PCIe Correctable Error Count in the SMART / Health Information Extended (Log Identifier C0h) SMART-14.
PCIERR-13	REPLAY_NUM Rollover Status (RRS)	
PCIERR-14	Bad DLLP Status (BDS)	
PCIERR-15	Bad TLP Status (BTS)	
PCIERR-16	Receiver Error Status (RES)	

5.3 Low Power Modes

Requirement ID	Description
LPWR-1	If Active State Power Management (ASPM) is supported, the default device state shall be disabled (see PCI Sig 3.1a or later).

5.4 PCIe Eye Capture

Requirement ID	Description
EYE-1	A utility shall be provided that will allow the user to capture the internal receive eye of the device to tune the signal integrity of the device to the target platform.

6 Reliability

6.1 Uncorrectable Bit Error Rate

Requirement ID	Description
UBER-1	The device shall support an Uncorrectable Bit Error Rate (UBER) of < 1 sector per 10^{17} bits read.

6.2 Power On/Off Requirements

6.2.1 Time to Ready and Shutdown Requirements

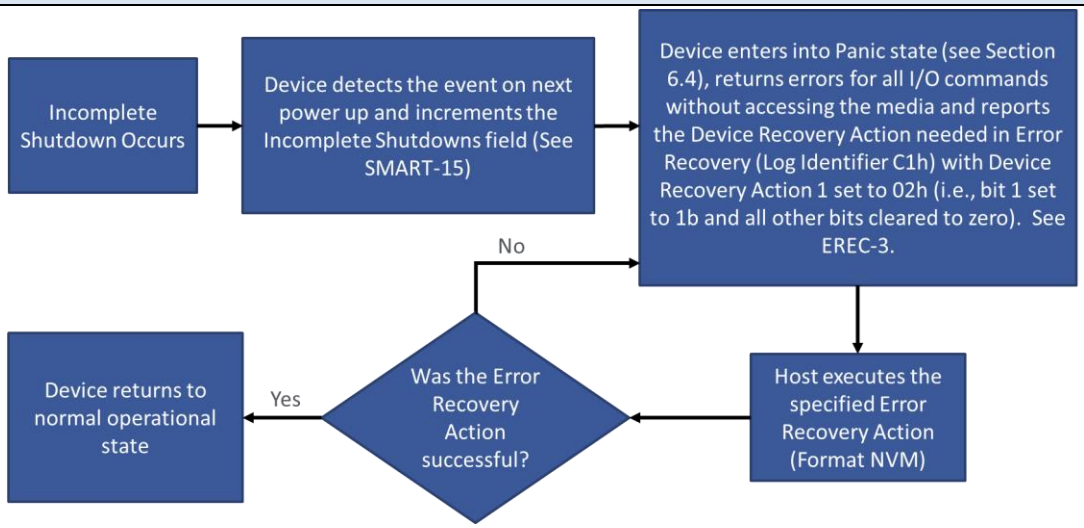
Requirement ID	Description
TTR-1	The device shall respond successfully to the Identify command within 1 second of CC.EN being set to 1b. Some models may allow a longer time (see Section 12 Device Profiles).
TTR-2	The device shall be able to process I/O commands with a successful completion within 20 seconds of CC.EN being set to 1b. Some models may allow a longer time (see Section 12 Device Profiles).
TTR-3	The device shall successfully service Admin commands as soon as CSTS.RDY = 1.
TTR-4	The device shall keep CSTS.RDY = 0 until the device is able to service Admin commands. Some models may allow a longer time (see Section 12 Device Profiles).
TTR-5	The Shutdown Notification completion (CSTS.SHST) shall be received within 10 seconds of setting SHN bit.
TTR-6	The device Controller shall support the CC.SHN (Normal and Abrupt Shutdown Notifications) at a minimum.
TTR-7	When safe shutdown is completed successfully, the device shall not enter a rebuild/recovery mode on the next power on.
TTR-8	Shutdown Notification shall trigger flushing of all content within the device's internal (SRAM/ DRAM) cache(s) (if one is present).
TTR-9	Obsolete. See NVMe-CFG-3 (support reporting of CSTS.CFS).
TTR-10	Obsolete. Moved to PLP-1 (full power-loss protection).
TTR-11	Obsolete. Moved to PLP-2 (protection health check).
TTR-12	Metadata rebuild due to any reason (e.g., unexpected power loss) shall not exceed 120 seconds and the device shall meet its latency requirements after this.
TTR-13	Obsolete. Moved to PLP-3 (protection health check interval).
TTR-14	Obsolete. Moved to PLP-4 (protection health check energy).
TTR-15	In case of a successful Normal shutdown operation (CC.SHN = 1 set by the NVMe device driver), no data loss is tolerated even if PLP has failed.
TTR-16	An Abrupt shutdown event shall not make the device non-functional under any conditions.
TTR-17	Obsolete. Moved to PLP-5 (false detection of power loss protection).
TTR-18	Obsolete. Moved to PLP-6 (capacitor holdup energy margin).
TTR-19	When the CC.SHN register is written to notify the device to shutdown it shall not be assumed that power will be lost even after CC.EN is cleared to 0. Under these

Requirement ID	Description
	conditions the device shall continue to function properly based on the NVMe and PCIe Specifications.
TTR-20	The device shall be able to successfully respond to Configuration space accesses within 1s from de-assertion of PERST#.

6.2.2 Incomplete/Unsuccessful Shutdown

An incomplete/unsuccessful shutdown is a Normal or Abrupt power down that did not complete 100% of the shutdown sequence for any reason that causes the device to be unable to guarantee data/metadata integrity (e.g., firmware hang/crash, capacitor failure, PLP circuit failure, etc.).

Requirement ID	Description
INCS-1	When the power-loss protection mechanism fails for any reason while power is applied, the device shall set Critical Warning bit 4 to 1b in the SMART / Health Information (Log Identifier 02h) and transition to the write behavior as defined in the EOL/PLP Failure Mode (Feature Identifier C2h).
INCS-2	The device shall incorporate a shutdown checksum or flag as the very last piece of data written to flash to detect incomplete shutdown.
INCS-3	This checksum in INCS-2 shall be used on power-up to confirm that the previous shutdown was 100% successful.
INCS-4	An incomplete shutdown shall result in an increase in the Incomplete Shutdowns field (SMART-15) in the SMART / Health Information Extended (Log Identifier C0h) and the SMART / Health Information (Log Identifier 02h) Critical Warning field shall have bit 3 set to 1b.
INCS-5	The device shall still support data eradication via the Sanitize command even if it is operating in Read Only Mode (ROM), and it shall support admin commands.
INCS-6	When the device increments the Incomplete Shutdowns field in SMART / Health Information Extended (Log Identifier C0h), it shall use the following recovery procedure at the next power up:

Requirement ID	Description
	

6.3 End to End Data Protection

Requirement ID	Description
E2E-1	All user data and metadata shall be protected using overlapping protection mechanisms throughout the entire read and write path in the device including all storage elements (registers, caches, SRAM, DRAM, NAND, etc.).
E2E-2	At least one bit of correction and 2 bits of detection is required for all memories. This shall be for all memories regardless of function.
E2E-3	The entire DRAM addressable space shall be protected with at least one-bit correction and 2 bits of detection scheme (SECDED). This includes but not limited to the following: • Flash translation layer (FTL). • Mapping tables (including metadata related to deallocated LBAs). • Journal entries. • Firmware scratch pad. • System variables. • Firmware code.
E2E-4	Silent data corruption shall not be tolerated under any circumstances.
E2E-5	The device shall include a mechanism to protect against returning the data from the wrong logical block address (LBA), including previous copies from same LBA, to the host. It is acceptable if the device stores additional/modified information to provide protection against returning wrong data to the host. Device shall perform host LBA integrity checking on all transfers to and from the media.

Requirement ID	Description
E2E-6	All device metadata, firmware, firmware variables, and other device system data shall be protected by at least a single bit detection scheme.
E2E-7	Any memory buffers that are utilized to accelerate data transfer (read-ahead buffers for example) shall follow the protection scheme outlined in E2E-5 (wrong logical block address protection).

6.4 Behavior on Firmware Crash, Panic or Assert

Requirement ID	Description
CRASH-1	After a firmware crash, panic or assert in which the device can no longer accept Write commands, the device shall allow read access only if it can guarantee data integrity.
CRASH-2	After a firmware crash, panic or assert, that is not a controller hardware failure, the device shall still support ADMIN commands including the ability to read any failure logs from the device to determine the nature of the failure.
CRASH-3	After the host performs the action specified in Device Recovery Action (see EREC-3 (Device Recovery Action)), the device shall allow full read and write access at full performance.
CRASH-4	If after a firmware crash, panic or assert there is the possibility of user data corruption, the Device Recovery Action shall require a Format (any SES type).

6.5 Annual Failure Rate (AFR)

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td>REL-1</td><td>The device shall meet an MTBF of 2.5 million hours (AFR of <= 0.35% per JEDEC JESD 218) under the following environmental conditions:
<table>
<tr><th>Specification</th><th>Environment</th><th>Requirement</th></tr>
<tr><td rowspan="2">Temperature</td><td>Operational</td><td>• 0˚C to 50˚C (32˚F to 122˚F)
• < 20°C (68˚F) per hour gradient</td></tr>
<tr><td>Non-Operational</td><td>• -40˚C to 70˚C (-40˚F to 158˚F)
• < 30˚C (86˚F) per hour gradient</td></tr>
<tr><td>Humidity</td><td>Operational</td><td>• 10% to 90% non-condensing
• Yearly weighted average: < 80% RH
  ○ 90% of year: < 80%
  ○ 10% of year: 80% to 90%
• Maximum dewpoint: 29.4°C (85°F)</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td></td><td>
<table>
<tr><td></td><td>Non-Operational</td><td>• 5% to 95% non-condensing
• 38°C (100.4°F) maximum wet bulb temperature</td></tr>
</table>
</td></tr>
<tr><td>REL-2</td><td>The device shall meet an MTBF of 2.0 million hours (AFR of <= 0.44% per JEDEC JESD 218) under the following environmental conditions:
<table>
<tr><th>Specification</th><th>Environment</th><th>Requirement</th></tr>
<tr><td rowspan="2">Temperature</td><td>Operational</td><td>• 0˚C to 55˚C (32˚F to 131˚F)
• < 20°C (68˚F) per hour gradient</td></tr>
<tr><td>Non-operational</td><td>• -40˚C to 70˚C (-40˚F to 158˚F)
• < 30˚C (86˚F) per hour gradient</td></tr>
<tr><td rowspan="2">Humidity</td><td>Operational</td><td>• 10% to 90% non-condensing
• Maximum dewpoint: 29.4°C (85°F)</td></tr>
<tr><td>Non-operational</td><td>• 5% to 95% non-condensing
• 38°C (100.4°F) maximum wet bulb temperature</td></tr>
</table>
</td></tr>
<tr><td>REL-3</td><td>Supplier shall provide the temperature conditions used to determine the MTBF.</td></tr>
<tr><td>REL-4</td><td>Supplier shall provide AFR de-rating curves for the Temperature range shown in requirement REL-2 for up to 70˚C (158˚F).</td></tr>
<tr><td>REL-5</td><td>The AFR targets in REL-1 (MTBF of 2.5 million hours) and REL-2 (MTBF of 2.0 million hours) shall be maintained up to a continuous reported composite temperature of 77˚C (170 °F) (WCTEMP) with less than 1% of the device power on time above WCTEMP.</td></tr>
</table>

6.6 Background Data Refresh

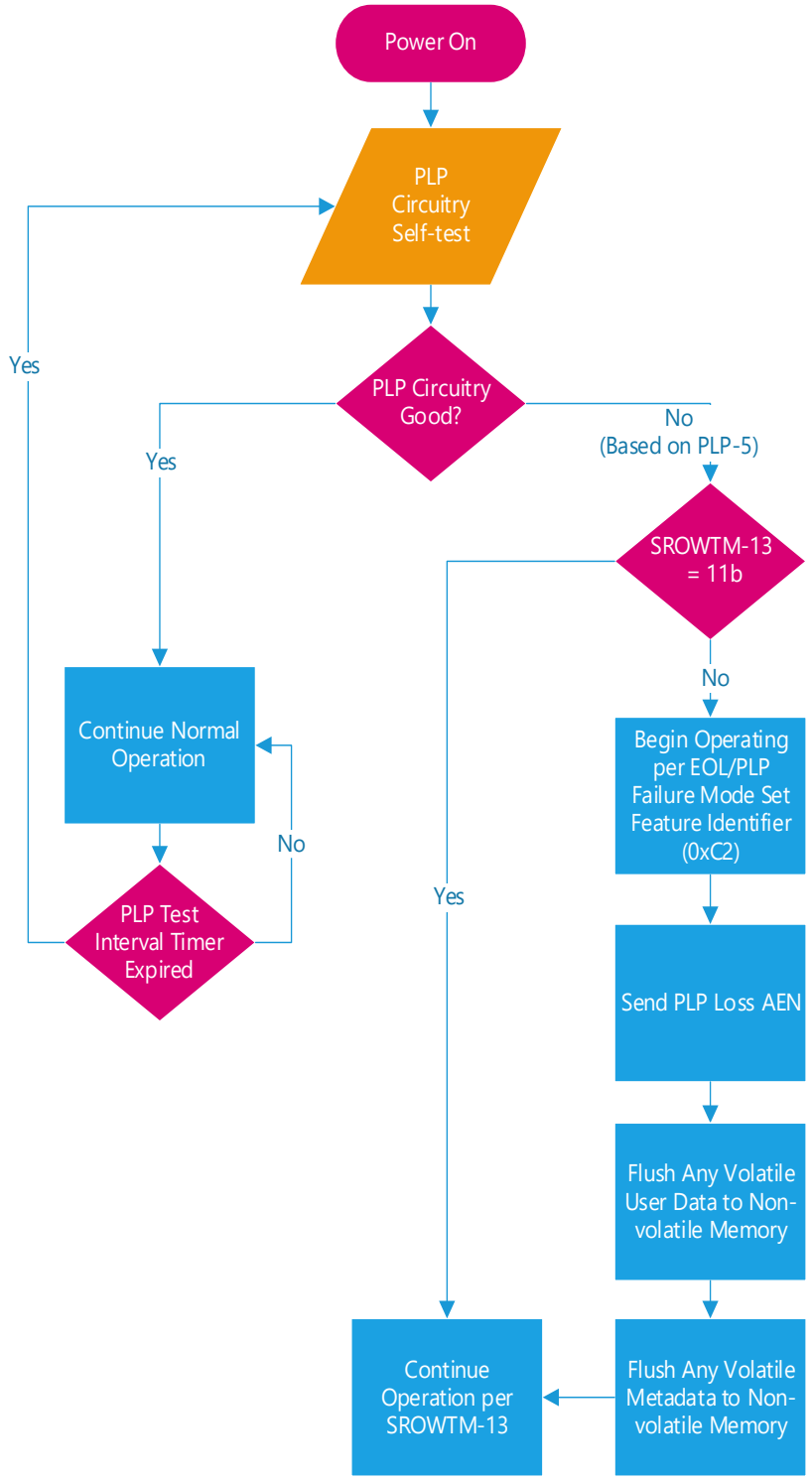
Requirement ID	Description
BKGND-1	The device shall support background data refresh while the device is powered on to ensure there is no data-loss due to power-on retention issues.
BKGND-2	The device shall be designed and tested to support the normal NAND operating temperature.
BKGND-3	Background data refresh shall cover the entire device and be designed to continuously run in the background and not just during idle periods.

6.7 Wear-leveling

Requirement ID	Description
WRL-1	The device shall utilize the entire Endurance Group media capacity range whenever the device needs to wear-level a block. The device shall not restrict the wear-leveling range to a subset of the Endurance Group media capacity unless otherwise specified. If the device does not support Endurance Groups, it shall wear-level across the entire physical media.

6.8 Power Loss Protection

Requirement ID	Description
PLP-1	The device shall support full power-loss protection for all acknowledged data and metadata.
PLP-2	The Power-loss protection health check shall not impact I/O latency and performance.
PLP-3	Power-loss protection health check shall be performed by the device at power-on (prior to accepting any writes) and at least once every time interval as specified in SPLPI-13 (PLP Health Check Interval).
PLP-4	While performing the power-loss protection health check, the device shall still have enough energy to be able to handle an abrupt power loss properly.
PLP-5	The firmware/hardware algorithm shall deploy safeguards to prevent a false detection of power loss protection failure. Example of a false detection would be a glitch in any of the power loss circuitry readings which would cause a transient event to trigger a false power loss protection failure when the power loss protection hardware is healthy.
PLP-6	The device shall implement a power-loss protection (PLP) health check which can detect the capacitor holdup energy margin reported in the Capacitor Health field specified in SMART-19 (Capacitor Health) in SMART / Heath Information Extended (Log Identifier C0h). The PLP health check shall not just check for open/short capacitor conditions but shall measure the true available margin energy.
PLP-7	The factory default PLP Health Check Interval shall be 15 minutes (see GPLPI-2 (PLP Health Check Interval)).
PLP-8	The default PLP Health Check Interval of 15 minutes (see GPLPI-2 (PLP Health Check Interval)) shall not cause the device to violate REL-1 or REL-2.
PLP-9	The device shall use the following workflow for the PLP Health Check (see INCS-1 (power-loss protection mechanism fails) for PLP AEN details):

Requirement ID	Description
	

6.9 Device Limits

Requirement ID	Description
DEVLMT-1	The device shall not have any architectural restrictions on the number of times any of the following events can occur: • Firmware downloads and activation supported (see FWUP-2 (number of firmware downloads)); or • Changing password when taking/changing ownership via TCG; or • Crypto erase or block erase (format); or • Power cycles; or • Set/get features (including power state changes); or • Log page or debug log retrievals. This is not meant to be a test until fail requirement.

7 Endurance

7.1 Endurance Data

Requirement ID	Description
ENDUD-1	The device documentation shall include the number of physical bytes able to be written to the device assuming a write amplification of 1. The units should be gigabytes (10^9 bytes).
ENDUD-2	The preconditioning steps to test device performance at end of life are: • 50/50 Read/Write workload (by number of I/Os). • 4kiB Read accesses aligned to 4kiB boundaries. • 128kiB Write accesses aligned to 128kiB boundaries. • Random pattern of Read addresses. • Sequential pattern of Write addresses. • 100% active range. • 80% full device (80% data, 20% free space). • 0% compressible data. • Ambient temperature 35°C (95°F). • Short-stroked device if capacity is 2TB or greater (see EOL-2 (method to "short stroke" the device)).

Requirement ID	Description
ENDUD-3	The Percentage Used in the SMART / Heath Information (Log Identifier 02h) shall track linearly with bytes written and at 100% it shall match the EOL value specified in ENDUD-1 (number of physical bytes able to be written).

7.2 Retention Conditions

Since there are several factors that impact the device endurance, the table below provides the requirements for the datacenter environment.

Requirement ID	Description
RETC-1	Non-Operational (Powered-off) device data retention time (end of life) shall be at least 1 month at 40°C (104°F). See Section 12 Device Profiles for specific retention requirements. Specific Devices Profiles may have longer data retention time requirements.
RETC-2	Operating (Powered-on) data retention shall be at least 7 years. For purposes of this requirement, the assumption is that the Terabytes Written (TBW) capability of the devices is used linearly over the lifetime. This requirement does not imply any specific warranty period.
RETC-3	The device shall not throttle its performance based on the endurance metric (endurance throttling).

7.3 Shelf Life

Requirement ID	Description
SLIFE-1	A new device may be kept as a datacenter spare and therefore shall be fully functional even if it sits on the shelf for at least 1 year at 20°C (68°F) before getting installed in the server. The device shall be new in box factory state.
SLIFE-2	Shelf-life shall be documented and provided to the customer.

7.4 End-of-Life (EOL)

Requirement ID	Description
EOL-1	Various types of samples are required for EOL testing: 1. Beginning of Life (Short stroked if required by EOL-2). 2. End of Life (Short stroked if required by EOL-2). 3. End of Life (Not short stroked if different than #2). EOL is defined as the Total Bytes Written (TBW) specification has been surpassed (see ENDUD-1 (number of physical bytes able to be written)) or the non-volatile

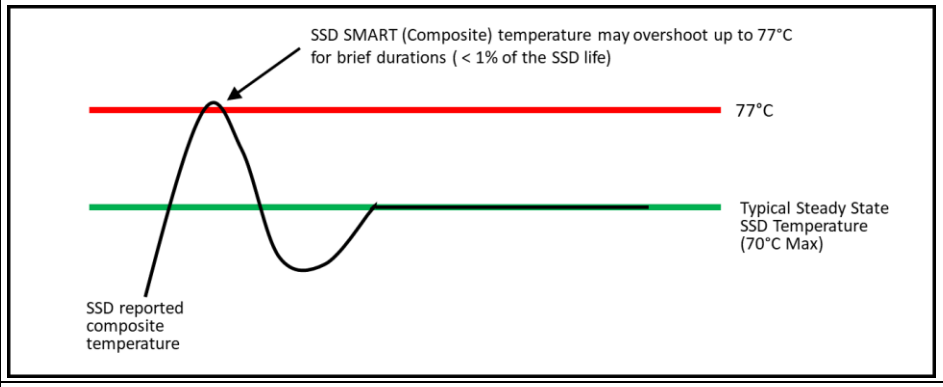
Requirement ID	Description
	media endurance limit (e.g., NAND cycling limit) has been reached; whichever is earliest.
EOL-2	On 2 TB or larger devices, there shall be a method to "short stroke" the device. Media reserved for background operations shall be proportionally adjusted. Short Stroke firmware capacity shall be reduced to a range of 10%-3% of full capacity. Short stroke shall cover all the channels/dies/planes while maintaining the native performance of the device.
EOL-3	Upon reaching 100% of specified device endurance, the device shall notify the host with an AEN.
EOL-4	The device shall continue to operate in a read/write mode until the conditions in EOL-5 are reached.
EOL-5	The device shall switch to Read Only Mode (ROM) when the available spares field in the SMART / Health Information (Log Identifier 02h) reaches 0%. A value of 0% represents the device state where there is an insufficient number of spare blocks to support Host writes. After the device switches to read-only mode, bit 2 and bit 3 of the Critical Warning field of Section 5.14.1.2 SMART Attributes in the NVMe specification version 1.4b shall be set to 1b. The device shall set bit 0 of the Critical Warning field to 1b in the SMART Attributes specified in Section 5.14.1.2 of the NVMe specification version 1.4b and generate a Critical Warning async notification (AEN) when the available spares value falls below the available spare threshold.

8 Thermal

8.1 Data Center Altitude

Requirement ID	Description
THERM-1	Support for devices in data centers located at an altitude of up to 10,000 feet above sea level is required (e.g., thermals, cosmic rays, etc.). When considering thermals, there shall be no de-rating up to 6,000 feet above sea level. Above 6,000 feet the derating shall be 0.9°C (1.6°F)/1000ft.
THERM-2	A thermal study with each platform is required and shall be provided to the customer. The thermal design shall be validated up to 35°C (95°F) ambient temperature for the platform with a worst-case airflow of 1.5 meters per second at sea level.
THERM-3	The device shall operate normally with relative humidity to be between 10% and 90%.

8.2 Thermal Throttling

Requirement ID	Description
TTHROTTLE-1	The device shall implement a thermal throttling mechanism to protect the device in case of a failure or excursion that causes the device temperature to increase above its maximum specified temperature.
TTHROTTLE-2	When the device begins throttling performance due to an over temperature condition, an Asynchronous Event Request shall be completed with the Asynchronous Event Type field set to 001b (SMART / Health status) and the Asynchronous Event Information field set to 01h (Temperature Threshold). The device shall set bit 1 of the SMART / Health Information (Log Identifier 02h) Critical Warning field to 1b.
TTHROTTLE-3	Thermal throttling shall only engage under certain failure conditions such as excessive server ambient temperature or beyond the server's fan failure redundancy limit. The required behavior is illustrated below: 
TTHROTTLE-4	The firmware algorithm shall deploy safeguards to prevent a false activation of either thermal throttling or thermal shutdown. Example of a false activation would be a glitch in any of the sensor readings which would cause the composite temperature to reach the thermal throttling or thermal shutdown limit.
TTHROTTLE-5	A composite temperature of 77°C (170.6°F) shall be used for throttling.
TTHROTTLE-6	Thermal throttling shall not start based on the rate of temperature increase or slew rate.
TTHROTTLE-7	When the device is in the thermal throttling state and the temperature drops back down below 75°C (167°F), the device shall exit the thermal throttling state and an Asynchronous Event Request shall be completed with the Asynchronous Event Type field set to 001b (SMART / Health status) and the Asynchronous Event Information field set to 01h (Temperature Threshold). The device shall clear bit 1 of the SMART / Health Information (Log Identifier 02h) Critical Warning field to 0b.
TTHROTTLE-8	When the device reaches a critical temperature on any component it shall report a composite temperature of 85°C (185°F) and an Asynchronous Event Request shall be

Requirement ID	Description
	completed with the Asynchronous Event Type field set to 001b (SMART / Health status) and the Asynchronous Event Information field set to 01h (Temperature Threshold) before the device shuts off to protect itself.
TTHROTTLE-9	The device shall report a value of 015Eh (77°C) in the Warning Composite Temperature Threshold (WCTEMP) field of the Identify Controller Data structure.
TTHROTTLE-10	The device shall report a value of 0166h (85°C) in the Critical Composite Temperature Threshold (CCTEMP) field of the Identify Controller Data structure.
TTHROTTLE-11	If the Composite Temperature is greater than or equal to 77°C (170.6°F), the device shall update the Warning Composite Temperature Time field of the SMART / Health Information (Log Identifier 02h) once a minute until one of the following occurs: - The device Composite Temperature falls below 75°C (167°F), or - The device shuts down due to an over Critical Temperature condition, or - The device is power cycled.

8.3 Temperature Reporting

Requirement ID	Description
TRPT-1	The device shall expose the current raw sensor readings from all the sensors on the device.
TRPT-2	The device's device-to-device composite temperature variation shall be no more than +/- 1 degrees C. Two different devices shall not report a composite temperature greater than 2 degrees apart under the same environmental conditions, slot location, and workload.
TRPT-3	A single device's composite temperature shall not vary by more than +/-1°C (1.8°F) degrees once it is in a steady state under the same environmental conditions, slot location, and workload.
TRPT-4	The supplier shall provide to the customer the equation, settings, and thresholds used to calculate the composite temperature.

8.4 Thermal Shutdown

Requirement ID	Description
THRMS-1	The device shall shut down to protect itself against data loss or damage due to extreme temperatures. The shutdown temperature value shall be at a composite temperature of 85°C (185°F) or higher. Shutdown shall preserve all user data.

9 Form Factor Requirements

9.1 Generic Form Factor Requirements

Requirement ID	Description
GFF-1	The device shall be compliant to PCIe base specification 4.0 (or later).
GFF-2	Vendor shall provide a PCIe compliance report.
GFF-3	The device shall support lane reversal with all lanes connected or partially connected lanes (e.g., a x4 device shall support it for x4, x2, and x1 connections).
GFF-4	The device shall train to x1 when only one upstream port is available, to x2 when the upstream device provides only 2 lanes per device and to x4 when 4 lanes are present.
GFF-5	The device shall support hot swap on form factors that support hot swap.
GFF-6	A CAD file of each supported form factor shall be provided to the customer.

9.2 Power Consumption Measurement Methodology

Requirement ID	Description
PCM-1	The device Max Average Power (MAP) consumption for any workload shall not exceed the Maximum Power (MP) of the current NVMe Power State over a 1s window with a sampling rate of 2ms or better. The measurement duration shall be at least 15 minutes on a pre-conditioned device. This requirement defines the measurement methodology for Maximum Sustained Device Power as defined in SFF TA-1009 revision 3.0.
PCM-2	The device peak power for any workload shall be measured over a 100uS window with a sampling interval of 4uS or smaller. The measurement duration shall be at least 15 minutes on a pre-conditioned device.
PCM-3	For devices, whose max average power consumption is less than or equal to 25W the peak power shall be no more than 1.5x the max average power. For devices, whose max average power consumption is greater than 25W but less than or equal to 29W the peak power shall be no more than 37.5W. For devices, whose max average power consumption is greater than 29W, the peak power shall be no more than 1.3x the max average power. The max average power is defined in PCM-1.

9.3 Power Levels

Requirement ID	Description
PWR-1	The Power Management (Feature Identifier 02h) shall be supported and the NVMe Power State Descriptor table in Identify Controller Data Structure bytes 3071:2048 shall be filled out per the NVMe 1.4b specification.
PWR-2	Obsolete. See DCLP-9 (DSSD Power State Table Descriptors) and Section 4.8.10.2 DSSD Power State Descriptor.
PWR-3	The method of measurement for Maximum Average Power (MAP) is defined in PCM-1 (device max average power). MAP values are reported in the Maximum Power (MP) field of NVMe Power State Descriptors.
PWR-4	Power state entries above the maximum rated power envelope of the device shall not be populated in the NVMe Power State Descriptor table in Identify Controller Data Structure bytes 3071:2048.
PWR-5	The Set Features for Power Management (Feature Identifier 02h) with the SV bit 31 in Command Dword 10 shall be supported so that the power level can be set and will be saved across power cycles.
PWR-6	The device, regardless of form factor or capacity, shall have an idle power of 5 Watts or less per European regulation.
PWR-7	The device shall not consume more power than the PCI-SIG Slot Capabilities register Slot Power Limit regardless of other settings (e.g., DSSD Power State, NVMe Power State).

9.4 M.2 Form Factor Requirements

Requirement ID	Description
FFM2-1	The device shall adhere to the M.2 specification with a size of 22mm x 110mm.
FFM2-2	The bottom-side height shall not exceed 1.5mm.
FFM2-3	The top-side height shall not exceed 3.2mm. Some models may need to be thinner (see Section 12 Device Profiles).
FFM2-4	The device shall use an M key.
FFM2-5	The device shall support a minimum of PCIe Gen3 x4.
FFM2-6	The device shall support driving an activity LED through the connector via LED_1#.
FFM2-7	The LED should be lit solidly when power is applied and flashing when there is traffic going to the SSD.
FFM2-8	The device shall not use any pins that are defined in the m.2 specification for vendor unique functionality.
FFM2-9	The device shall support a protection scheme that protects against NAND block level failures.

FFM2-10	The protection scheme must also support NAND plane level failures without data or metadata loss.
FFM2-11	The Label shall be placed on the top side of the device.
FFM2-12	The device electricals shall follow the SMBus connection as described below and in PCI SIG M.2 3.0 1.2 or later.
FFM2-13	The device's SMBus protocol shall comply to version 3.1 (see *http://smbus.org/specs/SMBus_3_1_20180319.pdf*).
FFM2-14	The default max average power for M.2 devices shall not exceed 8.25W and the peak power shall comply to PCM-3 (peak power limit).

9.5 E1.S Form Factor Requirements

Requirement ID	Description
FFE1S-1	The device shall adhere to the latest revision of SFF-TA-1006.
FFE1S-2	At a minimum, the device shall support PCIe Gen3 x4.
FFE1S-3	The device shall support activity and error LEDs.
FFE1S-4	The activity LED shall be lit solidly when power is applied and flashing when there is traffic going to the device.
FFE1S-5	The device shall support a protection scheme that protects against NAND block level failures.
FFE1S-6	The protection scheme must also support NAND plane level failures without data or metadata loss.
FFE1S-7	The amber LED shall meet the requirements specified in SFF TA-1009. The functioning of the Amber LED shall be independent of the 12V, 3.3Vaux and the state of the PWRDIS pin.
FFE1S-8	The thermal performance of the 9.5mm, 15mm, and 25mm cases and their associated pressure drops shall be provided.
FFE1S-9	The PWRDIS pin shall be supported.
FFE1S-10	The SMBus electrical connections shall follow the "DC Specification For 3.3V Logic Signaling" as defined in SFF-TA-1009 revision 2.0. Including Vih1 with a max of 3.465V.
FFE1S-11	The device's SMBus protocol shall comply to version 3.1 (see *http://smbus.org/specs/SMBus_3_1_20180319.pdf*).
FFE1S-12	A x4 device shall only have a 1C connector.

9.6 E1.L Form Factor Requirements

Requirement ID	Description
FFE1L-1	The device shall adhere to the latest revision of SFF-TA-1007.
FFE1L-2	At a minimum, the device shall support PCIe Gen3 x4.
FFE1L-3	The device shall support activity and error LEDs.
FFE1L-4	The activity LED shall be lit solidly when power is applied and flashing when there is traffic going to the device.
FFE1L-5	The device shall support a protection scheme that protects against NAND block level failures.
FFE1L-6	The protection scheme must also support NAND plane level failures without data or metadata loss.
FFE1L-7	The amber LED shall meet the requirements specified in SFF TA-1009. The functioning of the Amber LED shall be independent of the 12V, 3.3Vaux and the state of the PWRDIS pin.
FFE1L-8	The thermal performance of the 9.5mm and 18mm cases and their associated pressure drops shall be provided.
FFE1L-9	The PWRDIS pin shall be supported.
FFE1L-10	The SMBus electrical connections shall follow the "DC Specification For 3.3V Logic Signaling" as defined in SFF-TA-1009 revision 2.0. Including Vih1 with a max of 3.465V.
FFE1L-11	The device's SMBus protocol shall comply to version 3.1 (see *http://smbus.org/specs/SMBus_3_1_20180319.pdf*).
FFE1L-12	A x4 device shall only have a 1C connector.

9.7 E3 Form Factor Requirements

Requirement ID	Description
FFE3-1	The device shall adhere to the latest revision of SFF-TA-1008.
FFE3-2	The device shall adhere to the latest revision of SFF-TA-1009.
FFE3-3	The device shall support Separate Refclk Independent SSC Architecture (SRIS) requirements of the PCIe Express Base Specification Revision 3.0 and its ECNs. Devices shall support SRIS detection as described in SFF-TA-1009. The SRIS support shall include a clock tolerance of a 5600-ppm difference for separate reference clocks.

9.8 SFF-8639 (U.2/U.3) Form Factor Requirements

Requirement ID	Description
FF8639-1	The device shall adhere to the latest revisions of SFF-TA-1001 and PCI Express SFF-8639 Module Specification.
FF8639-2	The device shall support Separate Refclk Independent SSC Architecture (SRIS) requirements of the PCIe Express Base Specification Revision 3.0 and its ECNs. Devices shall support SRIS detection as described in SFF-8639 Module Specification Rev 4.0 or later. The SRIS support shall include a clock tolerance of a 5600-ppm difference for separate reference clocks.
FF8639-3	Activity LED All 2.5 SFF NVMe devices shall support the activity LED function on Pin 11 of the SFF-8639 connector. Pin 11 shall assert (see FF8639-4 (Active State)) while processing a command, the device shall drive the activity LED signal to an asserted state for 50ms and then to a negated state for 50ms (i.e., The LED is usually off, but flashes on and off while commands are processed). The activity cycle time frame for the blinking (assertion time + de-assertion time) shall be 500 milliseconds. If there is no activity the signal shall be de-asserted.
FF8639-4	Active State (pin asserted and LED on) is defined as logic “0” (i.e., < 0.3 volts) on Pin 11. Inactive State is defined as logic “1” (i.e., > 3.3 volts) on Pin 11. A device’s internally initiated background activity or maintenance routine that is not commanded to be performed from the host shall not cause the LED to be in the Active State (logic “0” level).

10 Out-of-Band Management Support

10.1 NVMe Basic Management Command (Appendix A) Requirements

Requirement ID	Description
SMBUS-1	The device shall support the NVMe Basic Management Command as defined in Appendix A of the NVMe Management Interface 1.1b specification:

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td></td><td>(https://nvmexpress.org/wp-content/uploads/NVM-Express-Management-Interface-1.1b-2020.10.05-Ratified.pdf).</td></tr>
<tr><td>SMBUS-2</td><td>SMBus Block Read protocol and Byte Read protocol shall be supported. As an override of the NVMe Management Interface 1.1b specification, SMBus Block Write protocol and Byte Write protocol shall also be supported.</td></tr>
<tr><td>SMBUS-3</td><td>The device shall implement the SMBus format as show in Section 10.2 - NVMe Basic Management Command (Appendix A) Data Format.</td></tr>
<tr><td>SMBUS-4</td><td>Unless otherwise noted, the default value for the Firmware Update Flags field (Byte 91) in the SMBus Data structure shall be set to FFh.</td></tr>
<tr><td>SMBUS-5</td><td>The Secure Boot Failure Feature Reporting Supported bit at offset 243 shall be supported and set to 1b.</td></tr>
<tr><td>SMBUS-6</td><td>When there is a secure boot failure the device shall report the failure with the following behavior:
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>7</td><td>Shall be set to 1b. See Command Code 242 (Secure Boot Failure Reporting) for definition.</td></tr>
<tr><td>6</td><td>Shall be set to 1b. See Command Code 242 (Secure Boot Failure Reporting) for definition.</td></tr>
<tr><td>5</td><td>May be set to 1b if OCP Recovery/ Platform Root-of-Trust for Recovery codes are supported. If this bit is set to 1b then a valid Recovery code shall be entered in byte 244.</td></tr>
<tr><td>4:0</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table></td></tr>
<tr><td>SMBUS-7</td><td>The device shall take no longer than the CAP.TO timeout value to produce stable SMBus output through the NVMe Simple Management Interface protocol (if supported).</td></tr>
<tr><td>SMBUS-8</td><td>If the device has set a non-zero value in the Panic ID field (Command Code 154, Byte 01h), then the device shall clear the Drive Functional field to 0b (Command Code 0, Byte 01h, Bit 5b).</td></tr>
<tr><td>SMBUS-9</td><td>The device shall generate the PEC values specified for each command code in the SMBus data format described in Section 10.2 - NVMe Basic Management Command (Appendix A) Data Format when the host issues a Block Read.</td></tr>
<tr><td>SMBUS-10</td><td>The device shall check the PEC value sent when the host issues a Block Write and only process the message if the PEC value matches the SMBus data format specified value described in Section 10.2 - NVMe Basic Management Command (Appendix A) Data Format. The device is encouraged to issue a NACK if the PEC value is not correct.</td></tr>
<tr><td>SMBUS-11</td><td>All data shall be returned in big-endian format unless otherwise noted.</td></tr>
</table>

10.2 NVMe Basic Management Command (Appendix A) Data Format

<table>
<tr><th>Command Code (Decimal)</th><th>Byte Offset (Decimal)</th><th>Description</th></tr>
<tr><td>0</td><td>0</td><td>Defined in NVM Express Management Interface 1.1b.</td></tr>
<tr><td>8</td><td>8</td><td>Defined in NVM Express Management Interface 1.1b.</td></tr>
<tr><td rowspan="3">32</td><td>32</td><td>Payload length of Command Code 32: This is the number of bytes until the PEC code. This shall be set to 10h.</td></tr>
<tr><td>48:33</td><td>GUID: This is a 16-byte Global Unique Identifier.
<table>
<tr><th>Byte Address</th><th>Value</th></tr>
<tr><td>33</td><td>73h</td></tr>
<tr><td>34</td><td>89h</td></tr>
<tr><td>35</td><td>20h</td></tr>
<tr><td>36</td><td>E5h</td></tr>
<tr><td>37</td><td>6Bh</td></tr>
<tr><td>38</td><td>EEh</td></tr>
<tr><td>39</td><td>42h</td></tr>
<tr><td>40</td><td>58h</td></tr>
<tr><td>41</td><td>9Ah</td></tr>
<tr><td>42</td><td>7Ah</td></tr>
<tr><td>43</td><td>CEh</td></tr>
<tr><td>44</td><td>BDh</td></tr>
<tr><td>45</td><td>B3h</td></tr>
<tr><td>46</td><td>5Fh</td></tr>
<tr><td>47</td><td>00h</td></tr>
<tr><td>48</td><td>85h</td></tr>
</table></td></tr>
<tr><td>49</td><td>PEC: An 8-bit CRC calculated over the SMBus address, command code, second SMBus address and returned data. The algorithm is defined in the SMBus Specification.</td></tr>
<tr><td rowspan="2">50</td><td>50</td><td>Payload length of Command Code 50: Indicated the number of additional bytes to read before encountering PEC. Shall be set to 26h.</td></tr>
<tr><td>51</td><td>Temperature Flags: This field reports the effect of temperature on the device’s performance.</td></tr>
</table>

<table>
<tr><th>Command Code (Decimal)</th><th>Byte Offset (Decimal)</th><th>Description</th></tr>
<tr><td></td><td></td><td>Temperature Throttling – Bit 7 is set to 1b when the device is throttling performance to prevent overheating. Clear to 0b when the device is not throttling.

Bits 6:0 shall be set to 1111111b.</td></tr>
<tr><td></td><td>52</td><td>Max Power Supported: This shall denote the Max Average Power (MAP) supported by this device rounded to the nearest watt. Some examples of how to use this is a 50W device is 32h, a 25W device is 19h, a 15W device is 0Fh, an 8.25W device is 8W which is 08h.</td></tr>
<tr><td></td><td>84:53</td><td>Configured NVMe Power State: This is a copy of the NVMe Power State Descriptor Data Structure of the currently configured Power State and is laid out in little endian format.</td></tr>
<tr><td></td><td>88:85</td><td>Total NVM Capacity: This field indicates the total usable NVM capacity in the NVM subsystem in GB in Hex (2048 GB in total capacity = 0000800h). This field is equivalent to the TNVMCAP field in the Identify Controller Data Structure.</td></tr>
<tr><td></td><td>89</td><td>PEC: An 8-bit CRC calculated over the SMBus address, command code, second SMBus address and returned data. The algorithm is defined in the SMBus Specification.</td></tr>
<tr><td>90</td><td>90</td><td>Payload length of Command Code 90: Indicates number of additional bytes to read before encountering PEC. Shall be set to 04h.</td></tr>
<tr><td></td><td>91</td><td>Firmware Update Flags: This field allows the host to control whether the current firmware allows new firmware images to be activated (see Section 11 Security for more information).
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>7</td><td>Written by host, read by device:
<table>
<tr><td>1b</td><td>Unlock Firmware Update
Device shall enable Firmware update</td></tr>
<tr><td>0b</td><td>Lock Firmware Update
Device shall block and error on Firmware download and activate commands</td></tr>
</table></td></tr>
<tr><td>6</td><td>Written by device, read by host:
<table>
<tr><td>1b</td><td>Firmware Update Unlocked
Device shall allow Firmware download and activate commands</td></tr>
<tr><td>0b</td><td>Firmware Update Locked</td></tr>
</table></td></tr>
</table></td></tr>
</table>

Command Code (Decimal)	Byte Offset (Decimal)	Description
		<table><tr><td></td><td><table><tr><td></td><td>Device shall block and error on Firmware download and activate commands</td></tr></table></td></tr><tr><td>5:0</td><td>Shall be set to 111111b.</td></tr></table> The device shall revert to the default Unlock/Lock state on the next power cycle. If the host attempts a Download Firmware command when the device is in the Firmware Update Locked state, the device shall fail the command with status Operation Denied.
	94:92	Reserved. Shall be cleared to zero.
	95	PEC: An 8-bit CRC calculated over the SMBus address, command code, second SMBus address and returned data. The algorithm is defined in the SMBus Specification.
96	96	Payload length of Command Code 96: Indicates number of additional bytes to read before encountering PEC. Shall be set to 38h.
	104:97	Firmware Version Number: This field shall indicate the activated firmware version that is running on the device after the firmware activation took place. The format of this field shall be as defined in field Firmware Revision (FR) Section 5.15.2.2 Identify Controller Data Structure of the NVMe specification version 1.4b.
	112:105	Security Version Number: This is the Security Version Number of the currently running firmware image. The supplier increments this number any time a firmware includes a fix for a security issue or critical firmware fix that customer agrees rollback prevention is required. This is a copy of SMART-22 – Security Version Number.
	152:113	Model Number: This shall be a copy of the data in the Model Number field of Identify Controller Data Structure (CNS 01h, byte offset 63:24).
	153	PEC: An 8-bit CRC calculated over the SMBus address, command code, second SMBus address and returned data. The algorithm is defined in SMBus Specification.
154	154	Payload length of Command Code 154: Indicates number of additional bytes to read before encountering PEC. Shall be set to 0Ah.
	155	Panic Rest Action: See EREC-2 (Panic Reset Action).
	156	Device Recovery Action 1: See EREC-3 (Device Recovery Action 1).
	164:157	Panic ID: See EREC-4 (Panic ID).
	165	Device Recovery Action 2: See EREC-14 (Device Recovery Action 2).

<table>
<tr><th>Command Code (Decimal)</th><th>Byte Offset (Decimal)</th><th>Description</th></tr>
<tr><td></td><td>166</td><td>PEC: An 8-bit CRC calculated over the SMBus address, command code, second SMBus address and returned data. The algorithm is defined in the SMBus Specification.</td></tr>
<tr><td>167</td><td>167</td><td>Payload length of Command Code 166: Indicates number of additional bytes to read before encountering PEC. Shall be set to 20h.</td></tr>
<tr><td></td><td>199:168</td><td>Panic Context Buffer: Vendor may record additional information about the Panic ID reported in Command Code 154. Any unused bytes shall be cleared to zero.</td></tr>
<tr><td></td><td>200</td><td>PEC: An 8-bit CRC calculated over the SMBus address, command code, second SMBus address and returned data. The algorithm is defined in the SMBus Specification.</td></tr>
<tr><td>201</td><td>241:201</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td rowspan="5">242</td><td>242</td><td>Payload length of Command Code 242: Indicates number of additional bytes to read before encountering PEC. Shall be set to 04h.</td></tr>
<tr><td>243</td><td>Secure Boot Failure Reporting:
<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>7</td><td>Secure Boot Failure Feature Reporting Supported:
When set to 1b the secure boot feature reporting is supported. When cleared to 0b the secure boot failure feature reporting is not supported.</td></tr>
<tr><td>6</td><td>Secure Boot Failure Status: When cleared to 0b there is no secure boot failure. When set to 1b there is a secure boot failure. This bit shall only be set if the Secure Boot Feature Supported bit is set to 1b and there is a secure boot failure.</td></tr>
<tr><td>5</td><td>OCP Recovery/ Platform Root-of-Trust for Recovery: When set to 1b, OCP Recovery/ Platform Root-of-Trust for Recovery codes are supported in byte 244. When cleared to 0b OCP Recovery/ Platform Root-of-Trust for Recovery codes are not supported and byte 244 shall be cleared to zero.</td></tr>
<tr><td>4:0</td><td>Reserved. Shall be cleared to zero.</td></tr>
</table></td></tr>
<tr><td>244</td><td>Recovery Code: OCP Recovery/Platform Root-of-Trust for Recovery code.</td></tr>
<tr><td>246:245</td><td>Reserved. Shall be cleared to zero.</td></tr>
<tr><td>247</td><td>PEC: An 8-bit CRC calculated over the SMBus address, command code, second SMBus address and returned data. The algorithm is defined in the SMBus Specification.</td></tr>
</table>

Command Code (Decimal)	Byte Offset (Decimal)	Description
248	248	Payload length of Command Code 248: Indicates number of additional bytes to read before encountering PEC. Shall be set to 06h.
	250:249	Data Format Version Number: Indicates the version of this mapping used in the device. Shall be set to 0004h.
	254:251	Reserved. Shall be cleared to zero.
	255	PEC: An 8-bit CRC calculated over the SMBus address, command code, second SMBus address and returned data. The algorithm is defined in the SMBus Specification.

10.3 NVMe-MI 1.1b Requirements

Requirement ID	Description
NVMe-MI-1	The device shall support NVMe Management Interface Specification version 1.1b or later.
NVMe-MI-2	The device shall support MCTP over SMBus.
NVMe-MI-3	The device shall support MCTP over PCIe VDM.
NVMe-MI-4	The device shall support SMBUS Fixed and Discoverable.
NVMe-MI-5	The device shall NACK any SMBus/I2C addresses not listed in the NVM Express Management Interface 1.1b.
NVMe-MI-6	The device shall support the GET UDID command.
NVMe-MI-7	The device shall have a VPD that is accessible over SMBus and supports IPMI Platform Management FRU Information. The VPD shall support all required elements that are defined in the NVM Express Management Interface Specification Revision 1.1b.
NVMe-MI-8	All mandatory NVMe Management Interface commands defined by the NVMe Management Interface Specification 1.1b shall be supported through MCTP over SMBus and MCTP over PCIe VDM.
NVMe-MI-9	All mandatory NVMe Admin commands defined by the NVMe Management Interface Specification 1.1b shall be supported through MCTP over SMBus and MCTP over PCIe VDM.
NVMe-MI-10	The following NVMe Admin commands shall be supported by MCTP over SMBus, even when the PCIe Links are not active, and MCTP over PCIe VDM. • Firmware Activate/Commit • Firmware Image Download • Sanitize • Security Send

Requirement ID	Description
	• Security Receive • Get Features • Set Features
NVMe-MI-11	The following Log Identifiers shall be supported by MCTP over SMBus, even when the PCIe Links are not active, and MCTP over PCIe VDM. • SMART / Health Information Log (02h) • Error Information Log (01h) • Firmware Slot Information (03h) • Device Self-test (06h) • Sanitize Status (81h)
NVMe-MI-12	The following Log Identifiers shall be supported by MCTP over PCIe VDM. • Persistent Event Log (0Dh) • Telemetry Host-Initiated (07h) • Telemetry Controller-Initiated (08h)
NVMe-MI-13	The following Feature Identifiers shall be supported by MCTP, even when the PCIe Links are not active, over SMBus and MCTP over PCIe VDM. • Power Management (02h) • Timestamp (0Eh) • Temperature Threshold (04h)

11 Security

11.1 Basic Security Requirements

Requirement ID	Description
SEC-1	The device shall support signed firmware binary update which is checked before firmware is activated. The device firmware shall be authenticated using cryptographic keys on every reboot and during firmware update.
SEC-2	The device shall support XTS-AES-256 or AES-256 hardware-based data encryption or better.
SEC-3	The device shall support anti-rollback protection for firmware. The anti-rollback protection shall be implemented with a security version which is different than the firmware version. If the security version of the firmware being activated is greater or equal to the current security version the firmware may be activated. If the

Requirement ID	Description
	security version of the firmware being activated is not equal or greater than the firmware being activated the firmware update shall fail.
SEC-4	The device shall support Crypto Erase (see NVMe-AD-5 (support Format NVM) and NVMe-AD-7 (support Sanitize)).
SEC-5	The device shall support Secure Boot (see Section 11.2 Secure Boot).
SEC-6	The device shall have a method of identifying a secure boot failure which does not require physical access to the device.
SEC-7	The device's cryptographic module shall be FIPS 140-3 capable per CMVP and shall follow the NIST 800-90 (A, B and C) specification.
SEC-8	The device shall implement only FIPS and NIST approved implementations and algorithms.
SEC-9	The device shall support Key revocation allowing a new key to be used for firmware validation on update. Preferred implementation is to allow for up to 3 key revocations.
SEC-10	The device shall support Opal v2.01 with mandatory support for the Locking feature, the Opal SSC feature, and the Datastore Table feature.
SEC-11	The device shall support Single User Mode feature set Version 1.00, revision 1.00.
SEC-12	The device shall support Configurable Namespace Locking (CNL) feature set Version 1.00, revision 1.00 with mandatory support for the Namespace Global Range Locking object. The Namespace Non-Global Range Locking object may be supported.
SEC-13	For some models, the IEEE 1667 silo will be required for eDrive support.
SEC-14	Supplier shall follow the Security Development Lifecycle (SDL), and provide a report with the following for each qualification-ready or production-ready firmware version: • The Threat Model. • Fuzz & Pen Tests. • Static Analysis. • Build Logs and Compiler Security Settings. Additional information about the SDL is available here: *https://www.microsoft.com/en-us/sdl/default.aspx*
SEC-15	Security audits, including firmware source code review, shall be provided to the customer. This will include Telemetry and debug logs, etc.
SEC-16	All signing keys shall be stored in a Hardware Security Module (HSM) that is either FIPS 140-2 Level 3 (or greater) certified or FIPS 140-3 Level 3 (or greater) certified.

Requirement ID	Description
SEC-17	Access/use of signing keys should be restricted to a small set of developers, following the principle of least privilege. Number of people with access and their corresponding roles shall be provided.
SEC-18	All debug ports shall be disabled before the device leaves the factory. Alternatively, the ports shall only be accessible in the field after a successful exchange of a challenge-response mechanism using an asymmetric crypto scheme (refer to NIST SP 800-63). The state shall be reset to inaccessible on any reset or power cycle.
SEC-19	All vendor unique commands, log pages or set features that are not explicitly defined in this specification or approved of in writing by the customer shall be disabled before the device leaves the factory. Alternatively, the commands/log pages/set features shall only be accessible in the field after a successful exchange of a challenge-response mechanism using an asymmetric crypto scheme (refer to NIST SP 800-63). The state shall be reset to inaccessible on any reset or power cycle.
SEC-20	Adversarial testing using red teams shall be conducted before qualification start. A report of items attempted, and results shall be provided to the customer.
SEC-21	Vendor shall provide timely notification to the customer of security issues and delivery of fixes to the customer: • Vendor shall document all security fixes with each firmware update. • Vendor shall notify end customer within 7 days of discovering security issues in the device hardware or firmware. • Notification of issues shall include the process and timeline of the vendor's commitment to fix the issue: ○ For privately disclosed vulnerabilities, the duration shall be no longer than 90 days. ○ For publicly disclosed vulnerabilities, the duration shall be no longer than 7 days. ○ Vendors shall notify the customers about the known CVEs and security issues and provide security-related updates before public announcement.
SEC-22	All Telemetry and debugging logs shall be human readable.
SEC-23	The device shall not include user data, passwords, keys and any secret or sensitive information in any Telemetry or debug logs.
SEC-24	All public keys shall be revocable.
SEC-25	Secure Boot Flow shall be based upon immutable firmware and Root Key. All mutable Key(s), Certificate(s) and/or firmware shall be cryptographically bound to the immutable firmware and Root Key.

Requirement ID	Description
SEC-26	Secure firmware update flow shall be immutable from exploitation and use immutable public keys.
SEC-27	If the Opal SP has been activated and the host has supplied a band password, the device shall use that password as an added source of randomness for its Deterministic Random Number Generator (DRNG). For example, the device may use KDF (PBKDF2) on the user provided band password to generate an output of equal length to the required band key length plus DRNG seed length. This output can then be utilized in parts, one part as Opal band key and the other part as an additional input to the DRNG. If the DRNG implementation does not allow taking additional input, then an alternate option would be to XOR this output (e.g., from PBKDF2 operation) with initial DRNG seed.
SEC-28	The device shall delete all keys from volatile memory as soon as they are no longer needed for operation during the current power on state.
SEC-29	The device shall only store host provided passwords, host provided keys, or any host provided secret information in non-volatile memory at any stage in an encrypted form. The encryption key for this protection shall not be stored in non-volatile memory. Device shall not store plaintext/cleartext secrets in any non-volatile memories.
SEC-30	The supplier must deliver key management and encryption flow diagram details (not source code) which includes: • Encryption algorithms and modes (e.g., RNG, wrapping function, key derivation function). • Key size and password length. • Any crypto information that is stored in nonvolatile memory. • Critical Security Parameters. Complete details about algorithm inputs (e.g., Initialization vector source and size, salt source and size, unique per device/product line/vendor)
SEC-31	The supplier must provide industry certification reports, if available, such as FIPS, NIST for device and device components such as TRNG, RNG, Crypto engine etc.
SEC-32	Log data and user data (data transferred from the Host in Write Commands) shall be stored on physically separate areas on the device. For example, in the system area and the user data area, respectively.
SEC-33	The vendor shall provide a comprehensive list of what is and what is not in the logs.
SEC-34	The device supplier shall develop and make available to the customer a software implementation of the device's crypto algorithm functionality for verification.
SEC-35	All telemetry and debug logs shall only be writable by device firmware after a successful secure boot.

Requirement ID	Description
SEC-36	The device shall set the default TryLimit to 5 for all Opal authorities. The value of the Tries column shall only be cleared to 0 after a power cycle.
SEC-37	TCG Opal devices shall support TCG Opal configuration (e.g., Level 0 discovery, ownership, etc.) through the SMBus interface via the NVME-MI Security Send and Security Receive NVMe Admin commands. This shall be supported even when the PCIe Links are not active (e.g., PERST# is asserted, secondary bus reset).
SEC-38	The device shall support TCG Feature Set: Block SID Authentication.

11.2 Secure Boot

The device shall support Secure Boot. There are two fundamental things to address for secure boot:

1. Secure boot rooted in hardware.
2. Core Root of Trust Measurement.

The vendor should follow the recommendations in the *TCG Publication for Hardware Requirements for a Device Identifier Composition Engine*. DICE coupled with *RIOT Core* and Source for *RIOT* can help implement Cryptographic Identity with implicit attestation.

Requirement ID	Description
SBT-1	The device shall comply with the *FIPS 186-4 Digital Signature Standard (DSS)* and the *Open Compute Security Project Publication for Secure Boot Requirements*.
SBT-2	For Core Root of Trust measurement, each device shall have a Cryptographic Device Identity.
SBT-3	The *TCG DICE standard*, or hardware based cryptographic identity shall be implemented.
SBT-4	The device should following the guidance in the *Commercial National Security Algorithm Suite* regarding quantum resistant algorithms and key sizes.
SBT-5	Secure boot flow shall be immutable for exploitation and use immutable public keys.

11.3 Distributed Management Task Force (DMTF) Security Protocol and Data Model (SPDM)

Requirement ID	Description
SPDM-1	The device shall support firmware measurement and identity authentication per the DMTF SPDM 1.1 or later specifications.

Requirement ID	Description
SPDM-2	The device shall support the PCI-SIG Component Measurement and Authentication (CMA) ECN.
SPDM-3	The device shall support SPDM-1 and SPDM-2 via MCTP over SMBus/I2C.
SPDM-4	The device shall support SPDM-1 and SPDM-2 via MCTP over PCIe VDM.
SPDM-5	The device shall support SPDM-1 and SPDM-2 via PCI-SIG Data Object Exchange (DOE).
SPDM-6	The leaf certificate shall include a Subject Alternative Name extension as defined in the CMA ECN.

11.4 Data Encryption and Eradication

Requirement ID	Description
DATAE-1	Obsolete. Duplicate of SEC-2 (XTS-AES-256 or AES-256 hardware-based data encryption).

12 Device Profiles

The following are device profiles. This section is intended to be firmware-based configuration settings configured by device suppliers when manufacturing a device. A device may be configured with a mix of A and/or B settings. Each customer shall provide their A/B preference for each configuration setting.

The following conventions are used for the Device Profile Table:

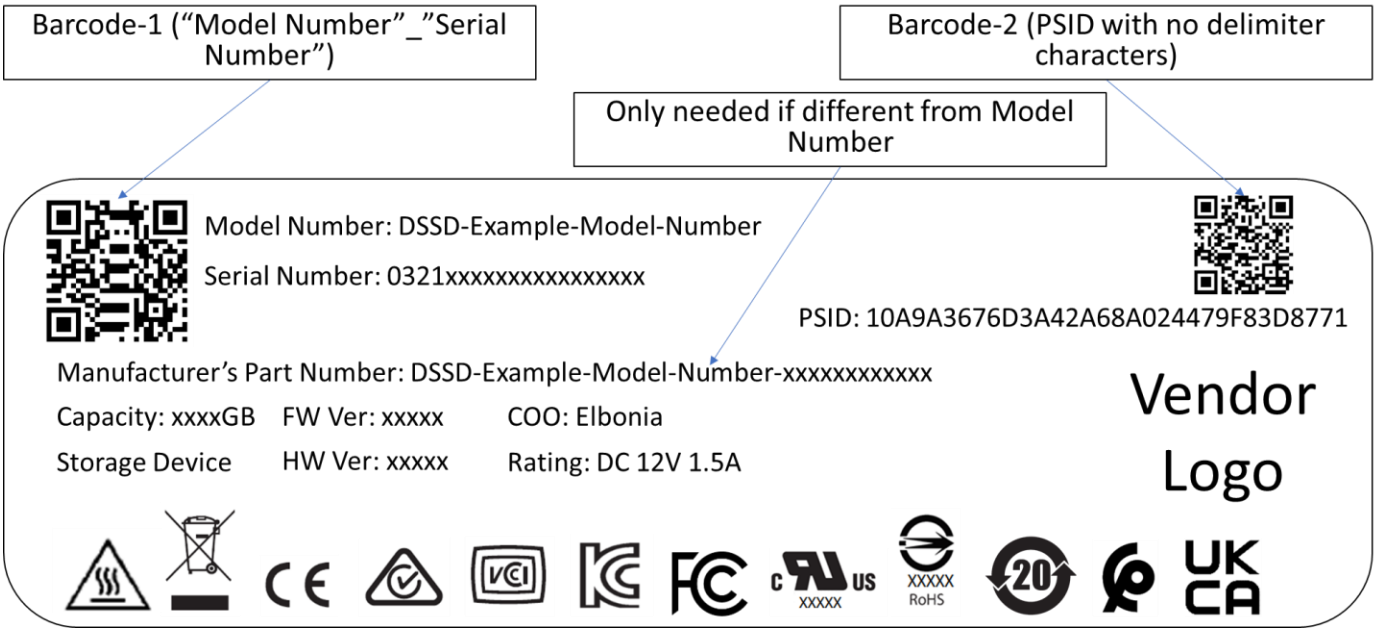
Convention	Definition
R	Required. This shall be supported.
O	Optional. This may be supported.
P	Prohibited. This shall not be supported.

Requirement ID	Description	Configuration Setting	
		A	B
DP-CFG-1	Factory Default Sector Size.	4096-byte	512-byte
DP-CFG-2	Number of Namespaces Supported.	NSM-4 (16 Namespaces)	NSM-5 (16 Namespaces per TB)
DP-CFG-3	Retention Time based on RETC-1 (data retention time).	1 Month	3 Months
DP-CFG-4	NVMe Basic Management Command Supported.	R	P
DP-CFG-5	Max M.2 top side height.	2.0mm	3.2mm

DP-CFG-6	EOL/PLP Failure Mode (Feature Identifier C2h).	Enabled	Disabled
DP-CFG-7	Write Uncorrectable command support.	O	R
DP-CFG-8	Time-to-Identify-Ready based on TTR-1.	<= 1 second	<= 10 seconds
DP-CFG-9	Time-to-I/O-Ready based on TTR-2.	<= 20 seconds	<= 10 seconds
DP-CFG-10	In addition to the requirements in TTR-4, the device shall keep CSTS.RDY = 0 until the device is able to service I/O commands successfully.	P	R

13 Labeling

The following sample label is meant to be used to refer to the label requirements in Section 13.1 – Label Requirements. It is not a model label and any markings on it are informative only. See the specifics in Section 13.1 Label Requirements for actual requirements:



13.1 Label Requirements

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td>LABL-1</td><td>The following fields are required information that shall be placed on the label:
<table>
<tr><th>Item</th><th>Format</th><th>Text Required</th><th>Barcode Required</th><th>Barcode Type</th></tr>
<tr><td>Barcode-1</td><td>‘Model Number’ ‘Underscore’ ‘Serial Number’ \n.</td><td>No</td><td>Yes</td><td>2d</td></tr>
<tr><td>Model Number</td><td>See LABL-11 (Model Number must match).</td><td>Yes</td><td>No</td><td>N/A</td></tr>
<tr><td>Serial Number</td><td>See LABL-12 (Serial Number must match), LABL-15 (certification logos), and LABL-17 (Serial Number format).</td><td>Yes</td><td>No</td><td>N/A</td></tr>
<tr><td>Manufacturer’s Part Number</td><td>Number used for ordering.</td><td>Yes, if different from Model Number</td><td>No</td><td>N/A</td></tr>
<tr><td>Capacity</td><td>Number of GB or TB.</td><td>Yes</td><td>No</td><td>N/A</td></tr>
<tr><td>STORAGE DEVICE</td><td>Text shall be “STORAGE DEVICE”.</td><td>Yes</td><td>No</td><td>N/A</td></tr>
<tr><td>PSID</td><td>TCG-OPAL Spec.</td><td>Yes</td><td>No</td><td>N/A</td></tr>
<tr><td>Barcode-2</td><td>‘PSID’ \n</td><td>No</td><td>Yes</td><td>2d</td></tr>
<tr><td>HW Revision</td><td></td><td>Yes</td><td>No</td><td>N/A</td></tr>
<tr><td>Firmware Name & Revision</td><td></td><td>No</td><td>No</td><td>N/A</td></tr>
<tr><td>Regulatory Mark</td><td colspan="2">See LABL-15 (certification logos).</td><td>No</td><td>N/A</td></tr>
<tr><td>Country Certification Numbers</td><td colspan="2">See LABL-15 (certification logos).</td><td>No</td><td>N/A</td></tr>
<tr><td>Certification Logos</td><td colspan="2">See LABL-15 (certification logos).</td><td>No</td><td>N/A</td></tr>
<tr><td>RoHS/ Green</td><td colspan="2">See LABL-15 (certification logos).</td><td>No</td><td>N/A</td></tr>
</table></td></tr>
</table>

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td>LABL-2</td><td>The Model Number on the shipping label shall match the Model Number used during qualification.</td></tr>
<tr><td>LABL-3</td><td>The minimum font size shall be 3 points and the typical size should be 6 points.</td></tr>
<tr><td>LABL-4</td><td>For the Capacity field, if there are space constraints, the manufacturer may remove “Capacity:” and just show “XXXGB” or “XXXTB”.</td></tr>
<tr><td>LABL-5</td><td>To distinguish Model Number and Serial Number, Barcode-1 shall have an underscore “_” between the Model Number portion and the Serial Number portion. Example:

Model Number: DSSD-Example-Model-Number
Serial Number: 0321xxxxxxxxxxxxxxxx

Barcod-1 Readout: DSSD-Example-Model-Number_0321xxxxxxxxxxxxxxxx</td></tr>
<tr><td>LABL-6</td><td>There shall be a line with the text “STORAGE DEVICE”.</td></tr>
<tr><td>LABL-7</td><td>The following fields are optional information that can be placed on the label at the discretion of the device maker. Placement is also at the device makers discretion if such information does not interfere with the mandatory information above. No additional barcode shall be present.
<table>
<tr><th>Item</th><th>Format</th><th>Text Required</th><th>Barcode Required</th><th>Barcode Type</th></tr>
<tr><td>Processor Code (BA)</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr>
<tr><td>Maker Logo</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr>
<tr><td>Rated Voltage & Current</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr>
<tr><td>Production Date</td><td>DDMMYYYY: DD (Date), MM (Month), YYYY (Year)</td><td>Optional</td><td>No</td><td>N/A</td></tr>
<tr><td>Weekly Code</td><td>YYWW: YY (Year), WW (Week)</td><td>Optional</td><td>No</td><td>N/A</td></tr>
<tr><td>Warranty VOID IF REMOVED</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr>
<tr><td>Makers Own Label Material Number</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr>
<tr><td>Website, Company Address</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr>
</table>
</td></tr>
</table>

Requirement ID	Description
	<table><tr><td>SSD</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr><tr><td>Product Series Name</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr><tr><td>SA: Value used within manufacturing</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr><tr><td>PBA: Physical Board Address (identifies the physical configuration of the device)</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr><tr><td>WWN: World Wide Number (unique for each device)</td><td></td><td>Optional</td><td>No</td><td>N/A</td></tr></table>
LABL-8	To ensure that datacenter operations personnel can quickly and easily identify devices that have been ticketed for field replacement, it is mandatory to have the proper identifying fields on the label(s), in the format specified below.
LABL-9	The label shall not degrade over the standard SSD lifetime under standard operating conditions.
LABL-10	For each formfactor, the label shall be placed as specified below: • M.2: the label shall be placed on the top side of the device as defined in the PCI-Sig M.2 formfactor specification. • E1.S: the label shall be placed on the Primary side of the device as defined in the SFF TA-1006 formfactor specification. • E1.L: the label shall be placed on the either the Primary or Secondary side of the device as defined in the SFF TA-1007 formfactor specification.
LABL-11	The Model Number in Barcode-1, the Model Number printed on the label and the Model Number returned in Identify Controller Data Structure (CNS 01h, byte offset 63:24) shall all match unless exempted in writing by the customer.
LABL-12	The Serial Number in Barcode-1, the Serial Number printed on the label and the Serial Number returned in Identify Controller Data Structure (CNS 01h, byte offset 23:04) shall all match.
LABL-13	This Hardware Revision printed on the label and returned by the NVMeCLI utility shall match.
LABL-14	All other electronically readable information shall also match their counterparts printed on the label.

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td>LABL-15</td><td>The following certification logos and their corresponding certifications are required:
<table>
<tr><th>Regulatory Mark/Text</th><th>Description</th></tr>
<tr><td>Regulatory Model Number</td><td>Unique regulatory Identifier.</td></tr>
<tr><td>Made in XXXX</td><td>Country of Origin.</td></tr>
<tr><td>Manufacturer or Brand name</td><td>Identification of the responsible party for placing the device into the market.</td></tr>
<tr><td>Address of the Manufacturer</td><td>Required for devices with the CE mark or UKCA mark.</td></tr>
<tr><td>Date of Manufacture</td><td>Not needed if embedded in the Serial Number.</td></tr>
<tr><td>Serial Number</td><td>Alpha-Numeric, 12-20 digits with first 4 digits indicating: Date of Manufacturing in Work Week and Year WWYY1234567890123456.</td></tr>
<tr><td>  </td><td>[Europe] Compliance with EU WEEE directive 2010/19/EU.</td></tr>
<tr><td>  </td><td>[Europe] Compliance with EU EMC directive 2014/30/EU and RoHS directive 2011/65/EU.</td></tr>
<tr><td>  </td><td>[Australia, New Zealand] Compliance with requirements of the relevant Australian ACMA Standards, under the Radiocommunications Act 1992 and the Telecommunications Act 1997.</td></tr>
<tr><td>  </td><td>[Japan] Compliance with Japan VCCI requirements.</td></tr>
<tr><td> 
XXXX-XX-XX</td><td>[Korea] Compliance with requirements of the Radio Research Laboratory Ministry of Information and Communication Republic of Korea.</td></tr>
<tr><td>CAN ICES-3(*)/NMB-3(*)</td><td>[Canada] Compliance with Canada standard ICES 003. Where * is either A or B.</td></tr>
<tr><td>  </td><td>[USA] Optional. Compliance with United States Federal Communications Commission requirement.</td></tr>
<tr><td> 
XXXXX</td><td>[USA] Compliance with UL standards and Canadian Safety Standards.</td></tr>
</table></td></tr>
</table>

Requirement ID	Description
	 [Taiwan] Compliance with Taiwan EMC and RoHS.  [China] Compliance with Chinese environmental requirements. Number inside the circle is usually 10 or 20.  [Morocco] Compliance with Moroccan EMC standards.  [United Kingdom] Mandatory after 1/1/2022. Guidance to UKCA marking.
LABL-16	If the surface of any component or casing will reach a temperature of 70°C (158°F) or greater the following warning logo shall be either printed on the label or placed separately on the device: 
LABL-17	The format of the Serial Number shall be WWYYSerialNumber with no leading spaces (e.g., WWYY1234567890123456).
LABL-18	The Model Number shall have no leading spaces.
LABL-19	Barcodes shall be printed using Datamatrix ECC200 or Model 2 QR code only.
LABL-20	QR codes shall use a minimum of ECC Level M (15%).
LABL-21	The density of the barcode shall be 10 mil or larger.
LABL-22	The label shall only be printed on Polyester or Plastic labels using a Wax/Resin ribbon.
LABL-23	The PSID shall be printed on the label in its direct 32-character alphanumeric representation without any ancillary delimiting characters (e.g., underscore, dash, backslash, forward slash, etc.) and shall exactly match the readout of Barcode-1.

14 Compliance

14.1 ROHS Compliance

Requirement ID	Description
ROHS-1	The Supplier shall provide component-level reporting on the use of listed materials by concentration (ppm) for all homogenous materials.

14.2 ESD Compliance

Requirement ID	Description
ESD-1	Device manufacturer needs to provide ESD immunity level (HBM- Human Body Model) measured in accordance with IEC-61000-4-2.

15 Shock and Vibration

Below are the shock and vibration specifications for storage devices:

Requirement ID	Description
SV-1	The non-operational shock requirement is 700G, half-sine, 0.5ms, total 6 shocks, along all three axes (+/-).
SV-2	The vibration requirement during operation is: 1.8G_{rms}, 5-500-5 Hz, Random Vibe, 20 min along all three axes.
SV-3	The vibration requirement during non-operation is: 3.13G_{rms}, 5-800-5 Hz, total 6 sweeps along all three axes, 20 minutes per sweep.
SV-4	Validation flow for Shock and Vibration: 1. UUT (Unit Under Test), test fixture should be visually inspected and ensured that everything is torqued or secured as needed. Pictures of test fixture with and w/o UUT should be provided. 2. Baseline performance of device should be gathered and used as a reference against post S&V data to ensure no performance impact incurred. 3. Once S&V testing is completed, repeat visual inspection to the UUT and test fixture to ensure no physical damage or performance impact has occurred to the UUT or test fixture. 4. Re-run stress test on the UUT in case of non-op test and provide data indicating no performance impact incurred to the unit

16 NVMe Linux CLI Plug-In Requirements

16.1 NVMe CLI Management Utility

The NVMeCLI utility (https://github.com/linux-nvme/nvme-cli) shall be used as one of the management utilities for NVMe devices.

Requirement ID	Description
UTIL-1	The SSD supplier must test their SSDs with this utility and ensure compatibility. The following is the minimum list of commands that need to be tested with NVMeCLI: • Format. • Secure erase. • FW update. • Controller reset to load FW. • Health status. • Log page reads including vendor log pages. • SMART status. • List devices. • Get/set features. • Namespace management. • Identify controller and namespace. • Effects log page.

16.2 NVMe CLI Plugin Requirements

The device supplier shall develop and provide a Linux NVMe CLI plugin that meets the following requirements:

Requirement ID	Description
UTIL-PI-1	A single, common plugin for all the supplier's NVMe-based products.
UTIL-PI-2	Vendor and additional log page decoding including into a human readable format and JSON output.
UTIL-PI-3	Obsolete.
UTIL-PI-4	The ability to pull crash dumps or FW logs (binary output is acceptable).
UTIL-PI-5	The plugin's subcommand nomenclature must adhere to Section 16.2.1 - NVMe CLI Plug-In Nomenclature/Functional Requirements and cannot change across versions.
UTIL-PI-6	The plugin shall use the existing NVMe CLI interface to access any vendor unique commands that are supported by the device.
UTIL-PI-7	If the NVMe CLI interface needs to transfer greater than 16 MB of data, the NVMe vendor unique Command shall have the ability to do multiple scatter/gather elements on the data buffer.

16.2.1 NVMe CLI Plug-In Nomenclature/Functional Requirements

The NVMe CLI plugin must meet the following naming and functional requirements:

Requirement ID	NVMe CLI Nomenclature	Purpose
UTIL-NM-1	vs-smart-add-log	Retrieve the SMART / Health Information Extended (Log Identifier C0h) from Section 4.8.4 - SMART / Health Information Extended (Log Identifier C0h). The plugin must use the exact same attribute names as indicated in that section.
UTIL-NM-2	vs-internal-log	Retrieves internal device telemetry/debug logging.
UTIL-NM-3	vs-fw-activate-history	Outputs the Firmware Activation History (Log Identifier C2h) in table format. See Section 16.2.2.1 NVMe CLI Plug-In FW Activation History Example Outputs for the table output format.
UTIL-NM-4	vs-drive-info	Outputs the following information: 1) Drive_HW_revision – Displays the current HW rev of the device. Any BOM or HW change must increment this version number. The value starts at 0 for pre-MP units and starts at 1.0 for MP units. The value increments by 0.1 for any HW changes in the pre-MP or MP stage. Qualification samples sent to Customer ODMs at the beginning of qualification is considered MP stage and needs to start at 1.0. This shall match the Hardware Revision printed on the label. 2) FTL_unit_size – Display FTL unit size. Units are in KB, so "4" means the FTL unit size is 4KB.
UTIL-NM-5	clear-pcie-correctable-errors	Calls Clear PCIe Correctable Error Counters (Feature Identifier C3h) Set Feature to clear the correctable PCIe error counters. See Section 4.12.7 Clear PCIe Correctable Error Counters (Feature Identifier C3h) Set Feature for more details.
UTIL-NM-6	clear-fw-activate-history	Calls Clear Firmware Update History (Feature Identifier C1h) Set Feature to clear the output of the "vs-fw-activate-history" and the Firmware Activation History (Log Identifier C2h). See Section 4.12.4 Clear Firmware Update History (Feature Identifier C1h) Set Feature for more details.
UTIL-NM-7	log-page-directory	The NVMe command that lists all the log pages and a description of their contents.
UTIL-NM-8	cloud-SSD-plugin-version	Prints version "1.0".
UTIL-NM-9	Help	Display this help.
UTIL-NM-10	vs-telemetry-controller-option	Obsolete.

16.2.2 NVMe CLI Plug-In FW Activation History Requirements

Requirement ID	Description
UTIL-FWHST-1	A table with entries that indicate the history of Firmware activation on the device.
UTIL-FWHST-2	Using the plugin command in UTIL-NM-4 (vs-drive-info) will retrieve the table.
UTIL-FWHST-3	Lists the last twenty recorded Firmware Activation History Entries as defined in FWHST-LOG-3. Oldest entries are on top.
UTIL-FWHST-4	When the device is first shipped from the factory, there are no entries recorded.
UTIL-FWHST-5	Obsolete. See FWHST-LOG-3 (Firmware Activation History Entry shall be recorded).
UTIL-FWHST-6	Redundant activation events shall not generate a new entry to prevent the scrolling out of useful information. An entry is redundant if it meets ALL the criteria below: 1. POH is within 1 minute from the last RECORDED entry. 2. Power cycle count is the same. 3. Current firmware is the same. 4. New FW activated is the same. 5. Slot number is the same. 6. Commit Action Type is the same. 7. Results are the same.
UTIL-FWHST-7	Firmware Activation History's output column headers shall follow the requirements below:

Requirement ID	Firmware Activation History Column Header	Purpose
UTIL-FWHST-8	Firmware Activation Counter	Increments every time a firmware activation is attempted no matter if the result is good or bad. Shall be cleared to zero at factory exit.
UTIL-FWHST-9	Power on Hour	Displays the POH of the SSD when the new firmware activation is completed, and the SSD is running with the new firmware. Accuracy needs to be down to the second.
UTIL-FWHST-10	Power Cycle Count	Display the power cycle count that the firmware activation occurred.
UTIL-FWHST-11	Previous Firmware	Displays the firmware that was running on the SSD before the firmware activation took place.
UTIL-FWHST-12	New Firmware Activated	Displays the activated firmware version that is running on the SSD after the firmware activation took place.
UTIL-FWHST-13	Slot Number	Displays the slot that the firmware is being activated from.

Requirement ID	Firmware Activation History Column Header	Purpose
UTIL-FWHST-14	Commit Action Type	Displays the Commit action type associated with the firmware activation event.
UTIL-FWHST-15	Result	Records the results of the firmware activation event. A passing event shall state a "Pass" for the result. A failing event shall state a "Failed" + the error code associated with the failure.

16.2.2.1 NVMe CLI Plug-In FW Activation History Example Outputs

FW Activation Examples:

Host FW download and activation events and initial states:
Initial State: Slot1 = 101
POH 1:00:00, PC 1, FW Commit CA = 011b Slot = 1 FW = 102
POH 2:00:00, PC 1, FW Commit CA = 001b Slot = 1 FW = 103
POH 3:00:00, PC 1, FW Commit CA = 001b Slot = 1 FW = 104
POH 4:00:00, PC 1, FW Commit CA = 001b Slot = 1 FW =105
Reset
POH 4:00:30 Reset complete, FW 105 activated
POH 5:00:00, PC 1, FW Commit CA = 011b Slot = 1 FW = 106
POH 6:00:00, PC 1, FW Commit CA = 001b Slot = 1 FW = 107
Power Cycle
POH 6:01:30 Power cycle complete, FW 107 activated
POH 7:00:00, PC 2, FW Commit CA = 001b Slot = 1 FW = 108

NVMe-CLI Plugin Output:

Firmware Activation Counter	Power on Hours	Power Cycle Count	Previous Firmware	New FW Activated	Slot Number	Commit Action Type	Result
1	1:00:00	1	101	102	1	011b	Pass
2	4:00:30	1	102	105	1	001b	Pass
3	5:00:00	1	105	106	1	011b	Pass
4	6:01:30	2	106	107	1	001b	Pass

Repeated Activation Events examples:

Host FW download and activation events and initial states:
Initial State: Slot1=101
POH 1:00:01, PC 1, FW Commit CA=011b Slot=1 FW=102, pass
POH 1:00:10, PC 1, FW Commit CA=0011b Slot=1 FW=102, fail reason #1
POH 1:00:30, PC 1, FW Commit CA=0011b Slot=1 FW=102, fail reason #1 (not recorded)

POH 1:01:15, PC 1, FW Commit CA=0011b Slot=1 FW=102, fail reason #1 (recorded as the time difference is greater than 1 minute from the last recorded event)
POH 1:01:25, PC 1, FW Commit CA=0011b Slot=1 FW=102, fail reason #2 (recorded as the failure reason changed)

NVMe-CLI Plugin Output:

Firmware Activation Counter	Power on Hours	Power Cycle Count	Previous Firmware	New FW Activated	Slot Number	Commit Action Type	Result
1	1:00:01	1	101	102	1	011b	Pass
2	1:00:10	1	102	102	1	011b	Fail #1
3	1:01:15	1	102	102	1	011b	Fail #1
4	1:01:25	1	102	102	1	011b	Fail #2

17 Revision History

Revision	Date	Release Notes
1.0	03/16/2020	Initial release based on feedback from the industry.
1.0a	06/26/2020	Errata.
2.0	05/18/2021	Additional major features and requirements including Latency Monitoring, Device Capabilities, Unsupported Requirements, Datacenter SSD Power States, Multiple Namespaces, Sanitize, NVMe-MI, Write Zeroes, Compare, Fused, Write Uncorrectable, Device Profiles, SPDM and additional security requirements, etc.

Appendix A – Facebook Specific Items

The following items apply specifically to devices delivered to Facebook.

A.1 Configuration Specifics

Requirement ID	Description
FB-CONF-1	Obsolete.
FB-CONF-2	IEEE 1667 shall not be supported. Devices shall not support Enable IEEE1667 Silo (Feature Identifier C4h) Set Feature or Enable IEEE1667 Silo (Feature Identifier C4h) Get Feature.
FB-CONF-3	For all form factors, SMBus byte 91 bit 6, Firmware Update Enabled bit shall be set to 1b by default from the factory.

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td>FB-CONF-4</td><td>Devices shall not support Error Injection (Feature Identifier C0h) Set Feature.</td></tr>
<tr><td>FB-CONF-5</td><td>Devices shall not support Error Injection (Feature Identifier C0h) Get Feature.</td></tr>
<tr><td>FB-CONF-6</td><td>Devices shall not support Error Recovery (Log Identifier C1h).</td></tr>
<tr><td>FB-CONF-7</td><td>All Telemetry and debugging logs can be either in binary or ASCII.</td></tr>
<tr><td>FB-CONF-8</td><td>The default power state shall conform to the following table:
<table>
<tr><th>Form Factor</th><th>Capacity</th><th>Default Power State Upon Factory Exit</th></tr>
<tr><td rowspan="2">E1.S</td><td><= 2TB</td><td>6 (12W)</td></tr>
<tr><td>>= 4TB</td><td>5 (14W)</td></tr>
</table></td></tr>
<tr><td>FB-CONF-9</td><td>Devices shall be configured to Configuration Setting A as shown in Section 12 - Device Profiles.</td></tr>
</table>

A.2 Performance Requirements

The following numbers are the Facebook performance targets for data storage SSD across all form factors. They are provided to serve as a guidance for SSD Vendors. Performance scripts can be found on GitHub at https://github.com/facebookincubator/FioSynth.

The targets are broken down into the following segments:

<table>
<tr><th>Requirement ID</th><th>Description</th></tr>
<tr><td>FB-PERF-1</td><td>FB-FIO Synth Flash Targets (for all capacities).</td></tr>
<tr><td>FB-PERF-2</td><td>fb-FIOSynthFlash TRIM Rate targets.</td></tr>
<tr><td>FB-PERF-3</td><td>IO.go benchmark target.</td></tr>
<tr><td>FB-PERF-4</td><td>Fileappend benchmark target.</td></tr>
<tr><td>FB-PERF-5</td><td>Sequential write bandwidth.</td></tr>
<tr><td>FB-PERF-6</td><td>Cache bench target.</td></tr>
<tr><td>FB-PERF-7</td><td>All targets shall be achieved by using “kyber” as the I/O scheduler.</td></tr>
<tr><td>FB-PERF-8</td><td>
<table>
<tr><th>Form Factor</th><th>Capacity</th><th>Max Average Power Consumption to Achieve All Performance Targets</th></tr>
<tr><td>M.2</td><td>2TB and smaller</td><td>8.5W</td></tr>
<tr><td rowspan="2">E1.S</td><td>2TB and smaller</td><td>10W</td></tr>
<tr><td>4TB</td><td>12W</td></tr>
</table>
The power measurement methodology is described in PCM-1 (device max average power), PCM-2 (device peak power), and PCM-3 (peak power limit).</td></tr>
</table>

A.2.1 Performance Targets for FB-FIO Synth Flash - HE_Flash_Short_TRIM_2H20 (for all capacities)

Syntax: fb-FioSynthFlash -w HE_Flash_Short_wTRIM_2H20 -d ALL - f HE_Flash_Short_wTRIM_2H20_results

Workload	Read MiB/s per TB	Write MiB/s per TB	TRIM BW per TB	P99 Read Latency	P99.99 Read Latency	P99.9999 Read Latency	Max Read Latency	P99.99 Write Latency	P99.9999 Write Latency	Max Write Latency
4K_L2R6DWPD_wTRIM	68 MiB/s	72 MiB/s	117 MiB/s	2,000 us	5,000 us	8,500 us	15,000 us	15,000 us	25,000 us	30,000 us
4K_L2R9DWPD_wTRIM	68 MiB/s	93 MiB/s	156 MiB/s	2,200 us	5,500 us	9,500 us	15,000 us	15,000 us	25,000 us	30,000 us
MyRocks_Heavy_wTRIM	210 MiB/s	101 MiB/s	22 MiB/s	2,000 us	5,000 us	8,500 us	15,000 us	10,000 us	20,000 us	25,000 us
Fleaf	320 MiB/s	87 MiB/s	89 MiB/s	3,000 us	6,000 us	10,000 us	15,000 us	20,000 us	25,000 us	30,000 us

A.2.2 Performance Targets for FB-FIO Synth Flash – Cache_2H20

Syntax: fb-FioSynthFlash -w Cache_2H20 -d ALL - f Cache_2H20

Workload	Read MiB/s per TB	Write MiB/s per TB	TRIM BW per TB	P99 Read Latency	P99.99 Read Latency	P99.9999 Read Latency	Max Read Latency	P99.99 Write Latency	P99.9999 Write Latency	Max Write Latency
B_Cache	164 MiB/s	96 MiB/s	0 MiB/s	2,000us	5,500 us	15,000 us	20,000 us	20,000 us	25,000 us	30,000 us

A.2.3 Performance Targets for FB-FIO Synth Flash – Search_2H20

Syntax: fb-FioSynthFlash -w Search_2H20 -d ALL - f Search_2H20

Workload	Read MiB/s per Node	Write MiB/s per Node	TRIM MiB/s per Node	P99 Read Latency	P99.99 Read Latency	P99.9999 Read Latency	Max Read Latency	P99.99 Write Latency	P99.9999 Write Latency	Max Write Latency
SearchLM_wTRIM	2,550 MiB/s	12 MiB/s	130 MiB/s	1,500 us	10,000 us	15,000 us	20,000 us	20,000 us	25,000 us	30,000 us

A.2.4 Performance Targets for FB-FIO Synth Flash – WSF

Syntax: fb-FioSynthFlash -w WSF -d ALL - f WSF

Workload	Read MiB/s per TB	Write MiB/s per TB	TRIM BW per TB	P99 Read Latency	P99.99 Read Latency	P99.9999 Read Latency	Max Read Latency	P99.99 Write Latency	P99.9999 Write Latency	Max Write Latency

wsf-t-heavy	4.4 MiB/s	35.5 MiB/s	76.9 MiB/s	400 us	2,800 us	3,000 us	3,300 us	3,700 us	4,200 us	5,000 us

A.3 Trim Rate Targets

- This test measures raw trim performance which no background I/O
- 64M trim >= 50GiB/s & <= 10ms P99 trim latency
- 3GB trim >= 500GiB/s & <= 10ms P99 trim latency

A.4 IO.go Benchmark Targets

- This test measures how long the file system is blocked from writing/overwriting a file while a different file is deleted
- Less than 4 file sizes total with latency outliers > 10ms
- No more than 2 latency outliers per file size
- No single latency outlier above 15ms

A.5 Fileappend Benchmark Targets

- This test measures how long the file system is blocked from appending to a file while a different file is deleted.
- No measurable stalls reported by this tool
- Max acceptable latency outlier is 10ms when deleting 1GiB or 2GiB file

A.6 Sequential Write Bandwidth

- Full device (all available user capacity, all namespaces) must be written/filled in 180 minutes or less
- Simple single-threaded sequential write FIO script to fill device

A.7 Cache bench target

- A benchmarking tool that's a supplement for FB FIO Synth Flash tool on measuring performance for cache applications. This is different than the "B Cache" workload in FB FIO Synth Flash.
- Two workloads need to be tested:
    - Tao Leader
    - Memcache
- The final allocator and throughput stats from the benchmark will be used to see if the targets are met.
- Send SSD latency versus time file to Facebook using one of the following methods:
    - Send the raw results log file
    - Run the "extract_latency.sh script and return the raw results log file, ".tsv" and ".png" files.

- Vendor NVMe CLI plug-in with “physical NAND bytes written” metric in the SMART / Health Information Extended (Log Identifier C0h) needs to be working to get the write amplification.

Workload	Get Rate	Set Rate	Read Latency (us)					Write Amp
			P50	P90	P99	P99.99	Max	
Tao Leader	87,000	16,000	100	300	800	3,000	12,000	1.3
Memcache WC	3,200	1,500	300	700	2,000	14,000	15,000	1.4
Workload	**Get Rate**	**Set Rate**	**Write Latency (us)**					**Write Amp**
			P50	**P90**	**P99**	**P99.99**	**P100**	
Tao Leader	87,000	16,000	30	50	100	700	8,000	1.3
Memcache WC	3,200	1,500	100	200	400	7,000	8,000	1.4

A.8 Max Latency Targets

This requirement aims to ensure the max latency for reads/writes are within certain bounds even as workload rate scales. We expect the workloads contained in this section to not exceed the max latency requirements defined here.

Workload	Max Read Latency (ms)	Max Write Latency (ms)
HE_Flash_Short_wTRIM_Sweep	65	85
Search_Sweep	65	85
Cache_Sweep	65	85

Appendix B – Microsoft Specific Items

The following items apply specifically to Microsoft.

B.1 Configuration Specifics

Requirement ID	Description
MS-CONF-1	E1.S form factor devices and M.2 form factor devices shall be formatted to 512 byte sectors from the factory.
MS-CONF-2	E1.L form factor devices shall be formatted to 4096 byte sectors from the factory.
MS-CONF-3	IEEE-1667 shall be supported.
MS-CONF-4	Obsolete.
MS-CONF-5	Obsolete.
MS-CONF-6	For all devices, SMBus byte 91 bit 6, Firmware Update Enabled bit shall be set to 1b (Firmware Update is Enabled) by default from the factory.

Appendix C – Latency Monitoring Feature Set Theory of Operation

C.1 Overview

Latency outliers are very undesirable in the data center. The goal of this feature is to allow production monitoring of SSD QOS outliers and to debug outlier issues in production. This feature will allow suppliers and hyperscale companies to clearly understand if the outliers are caused by the SSD or other components in the Host system. This feature enables SSD suppliers to effectively debug latency issues efficiently. This feature will enable predicting when latency outliers are growing and likely to impact hyperscale customers. This feature enables many use cases involving understanding and debugging latency issues at scale in a production environment.

C.2 Functional Operation

C.2.1 Latency Monitoring Feature Description

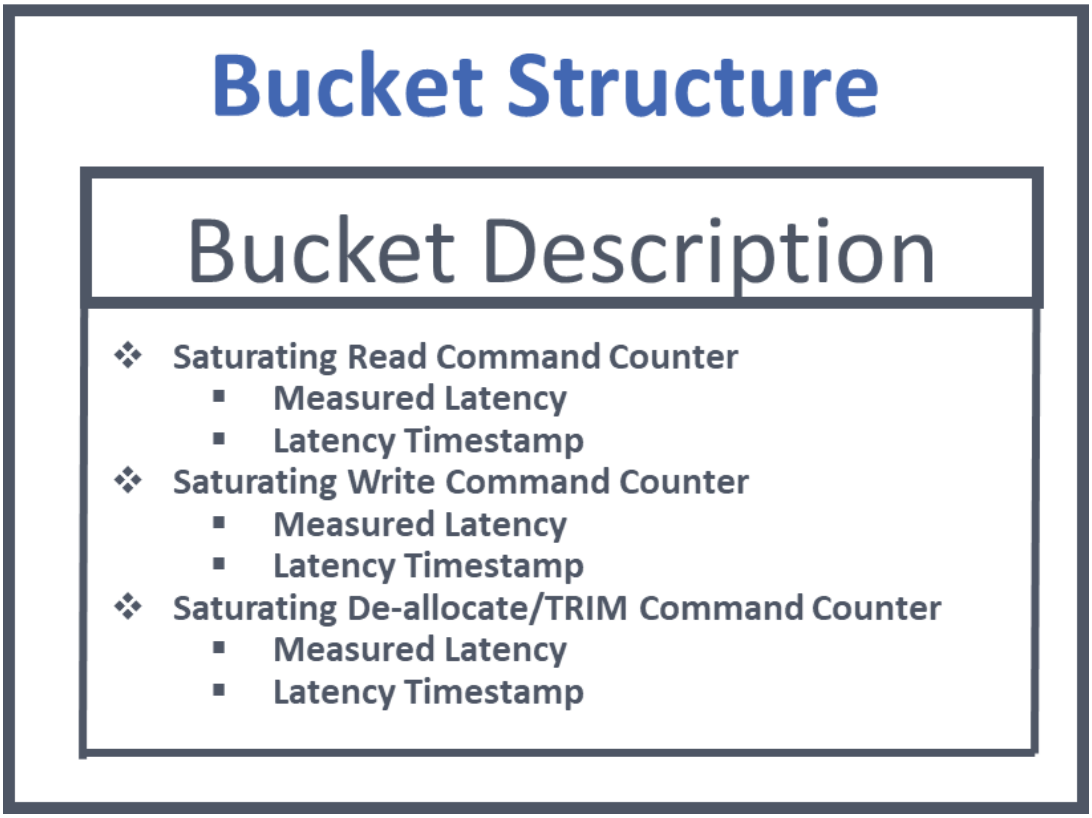
Below is the high-level theory of operation describing how the Latency Monitoring feature works. Latency of an individual command shall be measured from the time a controller fetches the command from the SQ to the time when controller writes the CQ entry for the SQ entry it fetched.

C.2.2 Bucket Description Overview

There are two types of buckets Active Buckets and Static Buckets. Active Buckets are buckets that are updated in real time. Static Buckets are buckets which are loaded with snapshots from the Active Buckets. This is a move from the Active Buckets to the Static Buckets, thus the old values in the Static Buckets are discarded. The Static Buckets allow hyperscale users to sample the Static Buckets over a fixed time period to gather statistics.

C.2.3 Active Bucket Description

The high-level concept is to create 4 real time buckets groups of active latency tracking command counters. Each bucket will count latency events which exceed a configured latency threshold. Below is a description of each bucket:



Each bucket contains the following:

- Saturating Read Command Counter with an associated Measured Latency and Latency Timestamp.
- Saturating Write Command Counter with an associated Measured Latency and Latency Timestamp.
- Saturating De-allocate/TRIM Command Counter with an associated Measured Latency and Latency Timestamp.

For clarity, the opcode to Counter mapping is below:

Bucket Counter	Opcode	Command
Read Command Counter	02h	Read
Write Command Counter	01h	Write
De-allocate/TRIM Command Counter	09h with Attribute – Deallocate (AD) = 1	Dataset Management

In addition to the command counters there is a Measured Latency data structure and a Latency Timestamp data structure associated with each command counter. The Measured Latency and Latency Timestamp have a direct relationship such that both are updated, or neither are updated. The Measured Latency and Latency Timestamp will be described later in this document.

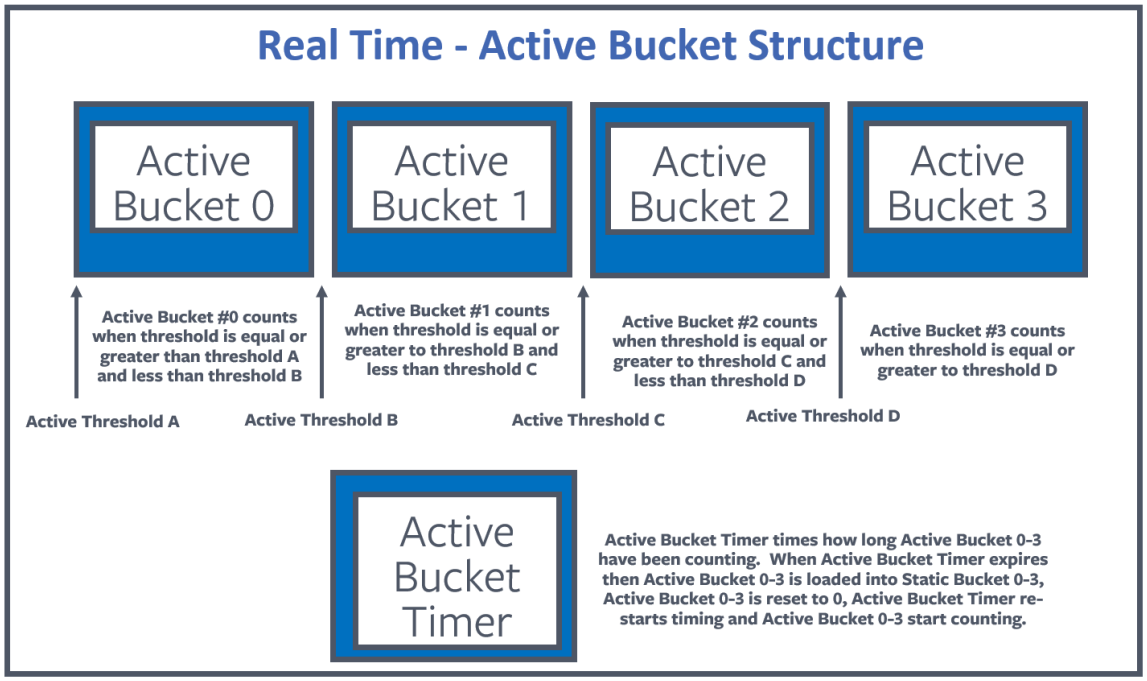
C.2.3.1 Active Command Counter Behavior

The active command counters count Read commands, Write commands, and De-Allocate/TRIM commands which exceed a configured latency threshold. These active command counters count until the command counter saturates or the Active Bucket Timer expires. Below is the behavior for each of these events:

- Active Command Counter Saturation:
  If the Command Counter saturates, the counter shall maintain the active value and not wrap.
- Active Bucket Timer Expiration:
  If the Active Bucket Timer Expires, then the following occurs:
  - The Active Bucket Command Counter values and associated information are moved into Static Bucket Command Counters, the Active Bucket Command Counters are then cleared to zero and re-start counting.
  - The active command counters shall count regardless of how the Active Latency Minimum Window is configured.

C.2.3.2 Active Bucket Thresholds

There are multiple Active Command Counter Buckets. The configured latency thresholds determine which Bucket the command counter shall be incremented in. Below is a picture showing how there are multiple Buckets and thresholds for each Bucket.



If the Read, Write, De-allocate/TRIM command completion time is below Threshold A then no counter is incremented. If the threshold is equal or greater than A and less than B, then the corresponding

command counter in Active Bucket 0 increments. If it is equal to or greater than B and less than threshold C, then the corresponding command counter in Active Bucket 1 increments. If it is equal to or greater than C and less than threshold D, then the corresponding command counter in Active Bucket 2 increments. If it is equal to or greater than threshold D, then the corresponding command counter in Active Bucket 3 increments. By following this process all latencies greater than threshold A are counted for Read, Write and De-Allocate/TRIM commands. When configuring the Latency Monitor Feature the thresholds shall always be configured such that Active Threshold A < Active Threshold B < Active Threshold C < Active Threshold D.

C.2.3.3 Active Bucket Timer Behavior

The Active Bucket Timer times how long the Bucket Command Counters have been counting. When the Active Bucket Timer is equal to the Active Bucket Timer Threshold then the following operations shall occur:

1. The following data is moved:
    a. Active Bucket Counters 0 - 3 is moved to Static Bucket Counters 0 - 3.
    b. Active Latency Timestamps are moved to Static Latency Timestamps.
    c. Active Measured Latencies are moved to Static Measured Latencies.
    d. Active Latency Timestamp Units are moved to Static Latency Timestamp Units.
2. The Active Bucket items shall then be updated as follows:
    a. Active Bucket Counters 0 - 3 are cleared to zero.
    b. Active Latency Timestamps are set to invalid (FFFF_FFFF_FFFF_FFFFh).
    c. Active Measured Latencies are cleared to zero.
    d. Active Latency Timestamp Units are cleared to zero.
    e. The Active Latency Minimum Window, if it is running, may be reset or continue to count.
    f. Active Bucket Timer is cleared to zero and starts the process of counting over.

When looking at the data structures in the Latency Monitor (Log Identifier C3h) it should be noted that the data in item #1 and #2 above is 16-byte aligned, and the data can be moved simply by doing a data move of the entire data structure.

C.2.3.4 Active Latency Timestamp Format

The format of the Latency Timestamp follows the Timestamp which is defined in NVMe 1.4b. The Latency Timestamp allows an understanding of where a latency excursion occurred in terms of time. The Latency Timestamp time reported shall be based on CQ completion.

If the device receives a Set Features with a Timestamp, then the device shall use this combined with the Power on Hours to determine the Latency Timestamp of when the latency event occurred. If the device receives multiple Set Features with a Timestamp the most recent Timestamp shall be used.

If the device does not receive a Set Features with a Timestamp, then the Latency Timestamp shall be generated based on Power on Hours.

If the device receives a Set Features with a Timestamp and then the device is powered off. When the device is powered on it shall use the most recent Timestamp it received even if this Timestamp was from before the device was powered off.

The Active Latency Timestamp Units shall be populated when the Latency Timestamp is updated to indicate if the Latency Timestamp used Timestamp with Power on Hours to generate the Latency Timestamp or if only Power on Hours were used to generate the Latency Timestamp.

C.2.3.5 Active Measured Latency and Active Latency Timestamp Updates

The Active Measured Latency is the latency measured from fetching the command in the SQ to updating the CQ. The Active Measured Latency data structure and the Active Latency Timestamp data structure shall be updated atomically. They shall not update independently. Each Command Counter (Read/Write/De-Allocate) has an Active Measured Latency and an Active Latency Timestamp structure associated with it. The Active Latency Configuration is used to configure this feature.

When the Latency Mode in the Active Latency Configuration is cleared to zero then the following behavior shall be followed:

- The Active Measured Latency and Active Latency Timestamp will be loaded the first time the command counter associated with it increments.
- The Active Measured Latency and Active Latency Timestamp will not be loaded again until the Active Measured Latency is reset.

If the Latency Mode is set to 0001h in the Active Latency Configuration, then every time the associated command counter is incremented the Active Measured Latency will report the largest measured latency based on the associated command counter. The Active Latency Timestamp will report the time when the largest latency occurred.

The Active Latency Minimum window also affects when the Active Measured Latency and the Active Latency Timestamp are updated. This is described in the section on Active Latency Minimum Window.

C.2.3.6 Active Latency Minimum Window

This affects both the Active Measured Latency and the Active Latency Timestamp. This defines the minimum time between updating the Active Measured Latency and the Active Latency Timestamp for a single Active Bucket/Counter combination. The feature is only used if the Active Latency Mode is set to 0001h.

If the Active Latency Minimum Window timer is running and the Measured Latency and Latency Timestamp have been updated, then the Latency Timestamp and Active Measured Latency will not be updated again until the Active Latency Minimum Window timer has expired. Below are some examples of this:

Example 1:
Assume:

- Active Latency Minimum Window of 5 seconds.
- Latency Mode is Configured for Largest Latency.
- Bucket 2 has a threshold range of 40ms to 400ms.

Time in seconds	Read Counter Bucket 2 Latency Event	Active Read Counter Bucket 2 Value	Actual Latency	Active Measured Latency	Latency Stamp	Comment
0	N	0	-	-	FFFF_FFFF_FFFF_FFFFh	Actual Latency and Active Measured latency are invalid.
0.5	Y	1	50ms	50ms	0.5 Seconds	First Latency Event. This starts the Active Latency Minimum window. New latency events will not be recorded until the 5 second Active Minimum Window expires at 5.5 seconds.
5.25	Y	2	100ms	50ms	0.5 Seconds	Measured Latency and Latency Timestamp is not updated due to Minimum Window is not expired; however, the Active Read Counter is updated.
6	Y	3	75ms	75ms	6 Seconds	Minimum Window is expired and 75ms is greater than previous number of 50ms, so the Active Measured Latency and Latency Timestamp are updated.

Example 2:
Assume:

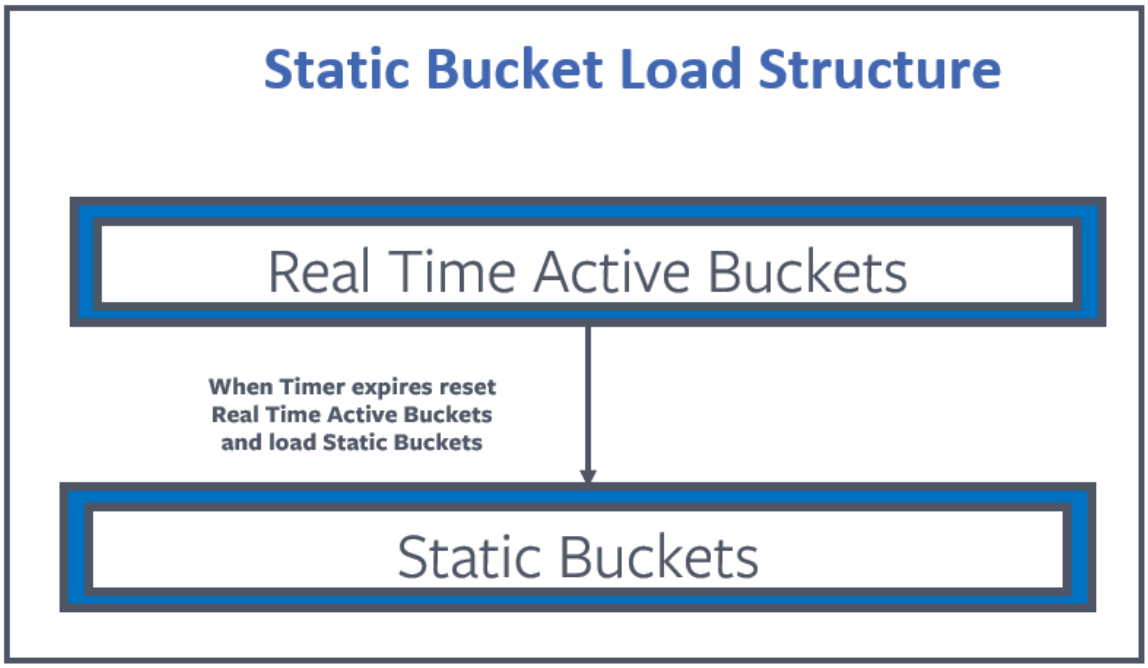
- Active Latency Minimum Window of 5 seconds.
- Latency Mode is Configured for First Latency Event.
- Bucket 2 has a threshold range of 40ms to 400ms.

Time in seconds	Read Counter Bucket 2 Latency Event	Active Read Counter Bucket 2 Value	Actual Latency	Active Measured Latency	Latency Stamp	Comment
0	N	0	-	-	FFFF_FFFF_FFFF_FFFFh	Actual Latency and Active Measured latency are invalid.
0.5	Y	1	150ms	150ms	0.5 Seconds	First Latency Event.
10	Y	2	200ms	150ms	0.5 Seconds	Since the device is in First Latency Event mode, no additional events are recorded.
15	Y	3	75ms	150ms	0.5 Seconds	Since the device is in First Latency Event mode, no additional events are recorded.

The Active Latency Minimum Window acts as a filter to ensure there are not a large number of events to update the Active Measured Latency and the Latency Timestamp. Thus, if the queue depth is 128 commands deep and there is a latency event, then there are not 128 updates to these data structures. Rather the first event is recorded, and the rest of the events are filtered out. It should be noted that the Active Latency Window is not enforced across power cycles. Thus, after a power cycle the Active Latency Window shall not start until a Bucket Counter is incremented.

C.2.3.7 Static Bucket Description

In addition to the real time buckets there are static buckets. When the Active Bucket Timer reaches its configured threshold (Active Bucket Timer Threshold) the active real time buckets shall be loaded into the static buckets and the active real time buckets shall be reset and start counting from 0. Below is a picture describing this:



C.3 Persistence Across Power Cycles

When either a safe or unsafe power transition happens the counters and associated Latency Monitoring Information shall be saved such that they can be restored on the next power up transition and continue from where the power cycle left off. There are additional details about this in the C.5 Latency Monitoring Feature - Challenging Event Handling section.

C.4 Debug Logs

The Latency Monitoring Feature can also enable debug logs to trigger. The Debug Log Trigger Enable configures which counters shall trigger a debug log the first time the Bucket/Counter combination is incremented. Only a single debug log shall be generated. Once a Latency Monitor debug log is generated, until the Latency Monitor debug log is discarded another Latency Monitor debug log cannot be generated. The Latency Monitor debug log shall be discarded using Set Features for the Latency Monitor or by reading the Latency Monitor Debug Log.

The Set Features for the Latency Monitor has two mechanisms for discarding the Debug Log. One method discards the debug log and resets the Latency Monitor feature to a new set of configured values based on the fields in Set Features. The other method discards the Debug Log and has no effect on any of the other features associated with the Latency Monitor Feature. Thus, the Latency Monitor Feature will keep running undisturbed when the Debug Log is discarded.

When the Latency Monitor debug log trigger event happens, the following data shall be captured: Debug Log Measured Latency, Debug Log Latency Timestamp, Debug Log Trigger Source, Debug Log Timestamp Units, Debug Log Pointer as well as internal information required to debug the issue to root cause.

C.5 Latency Monitoring Feature - Challenging Event Handling

C.5.1 Power Off/On When Latency Monitoring Feature is Enabled

When powering off, the Active Bucket Information (Counters, Measured Latency, Latency Timestamp, Active Bucket Timer) may be slightly off due to concerns with flushing data with unsafe power down. The Active Bucket Information shall maintain coherency compared to itself when flushing data with unsafe power down. When the device powers back on the Latency Monitoring Feature shall restore the Active/Static Bucket/Debug Information including loading the Active Bucket Counter. Once the restoration is complete then the device shall resume the Latency Monitor functionality. The device shall start capturing latency data within 2 minutes of power on. Thus, commands for the first 2 minutes may not be monitored. The Active Latency Window is not enforced across power cycles. Thus, after a power cycle the Active Latency Window shall not start until a Bucket Counter is incremented.

C.5.2 Power Off/On When Latency Monitoring Feature is Disabled

The Latency Monitoring Feature disabled/enabled state shall be persistent across power cycles. When the Latency Monitor Feature is enabled the associated Latency Monitor configuration information in the Latency Monitor Log shall persist across power cycles.

C.5.3 Moving Information from Active to Static Buckets

When moving data from the active buckets to static buckets it can be challenging to track all the information. The Active counter shall restart counting no longer than 3 seconds from moving data from Active to Static Counters.

C.5.4 Firmware Update

When activating new firmware, if the Latency Monitoring Feature is enabled, the firmware activation shall reset the Latency Monitoring Feature just as if a Set Features command to enable the feature was received. The Latency Monitoring Log shall start updating properly within 2 minutes of firmware activation completing. Thus, there are command latencies which could be missed after initially activating new firmware.

C.6 Configuring Latency Monitoring

Set Features is used to configure this feature. When Set Features is used to configure this feature all the data structures in both the Active and Static Buckets are reset. The debug log and related debug information is reset based on the Discard Debug Log Field.