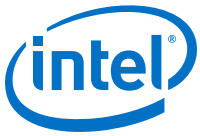
Intel® Solid State Drive D5-P4420 (U.2 Form Factor)

Product Specification

- Capacities:
  - 7.68TB[1]
- Performance[2]
  - Sequential Read/Write 128KB[3] QD128 Up to 3300/1000MB/s[4]
  - Random Read/Write 4KB[3] QD256: Up to 427K/46K
  - Average 4KB Random Latency Read/Write: 131µs/25µs[5]
- Intel® 64 layer QLC 3D NAND Flash Memory
- NVMe* Driver Support (64-bit)
  - Windows Server 2016, Windows Server 2012 R2, Win PE10*
  - RHEL* 7.4, 7.3, 7.2
  - CentOS* 7.5, 7.3
- Hardware-based AES-256 XTS Mode Encryption
- Compliance
  - NVM Express* 1.2.1
  - PCI Express* Base Specification Rev 3.1
- Certifications and Declarations
  - UL*, CE*, WEEE*, RCM*, BSMI*, KCC*, Maghreb*, VCCI*, ICES*, EFUP, RoHS*, Low Halogen
- Power
  - 12V Supply Rail & Optional 3.3V_{aux} Supply Rail for SMBus I^2C device
  - Enhanced power-loss data protection
  - Idle Power: 5W
  - Average write power:[6] 15W
  - Average read power:[7] 10W
- Endurance Rating
  - Sequential write up to 24.6 PBW[8]
  - Random write up to 5.6 PBW[9]
- Reliability
  - Uncorrectable Bit Error Rate (UBER): 1 sector per 10^{17} bits read
  - Mean Time Between Failure (MTBF): 2 million hours
- Temperature Specification
  - Operating Temperature (as reported by SMART): 0° C to 70° C with specified airflow
  - Non-Operating Temperature:[10] -55° C to 95° C
  - Temperature monitoring (In-band and out of band)
  - Thermal throttling at 70° C (SMART)
  - Thermal Shutdown at 80° C (SMART)
- Form Factor
  - U.2 15mm
- Approximate Drive Weight:
  - 125 grams
- Shock
  - Operating: 1000G (0.5ms)
  - Non-operating: 1000G (0.5ms)
- Vibration
  - Operating: 2.17 G_{RMS} (5-700 Hz)
  - Non-Operating: 3.13 G_{RMS} (5-800 Hz)
- Altitude
  - Operating: -1,000 to 10,000 ft.
  - Non-Operating: -1,000 to 40,000 ft.

Notes:

1. 1TB = 10^{12} Bytes
2. Performance specifications apply to both compressible and incompressible data
3. 4 KB = 4,096 bytes; 128 KB = 131,072 bytes
4. MB/s = 1,000,000 bytes/second
5. 50thpercentile 4K Read/Write QD1
6. The workload equates QD256/128KB Sequential Write. Average power is measured over a 100ms sample period
7. The workload equates QD256/128KB Sequential read. Average power is measured over a 100ms sample period
8. 1PB = 10^{15} Bytes
9. Workload: JESD219A
10. Please contact your Intel representative for details on the non-operating temperature range



Ordering Information

Contact your local Intel sales representative for ordering information.

Revision History

Revision	Description	Date
001	• Initial Release	January 2019
002	• Adjusted performance values in Table 35 and 37 • Added sequential write consistency values in Table 36 • Adjusted latency values in Table 38	May 2019
003	• Updated Quality of Service values in Table 40	May 2020

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order. This document contains information on products in the design phase of development.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase.

Performance results are based on testing as of October 16, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

Test and System Configuration: CPU: Intel(R) Xeon(R) Gold 6140 CPU 2.3GHz 24.75MB 140W 18 cores, BIOS: SE5C620.86B.00.01.0013.030920180427, CPU Sockets: 2, RAM Capacity: 32G, RAM Model: DDR4, RAM Stuffing: N/A, DIMM Slots Populated: Slots: 2, PCIe Attach: CPU (not PCH lane attach), Chipset: Intel® C610 chipset, Switch/ ReTimer Model/Vendor: Intel A2U44X25NVMEDK, NVMe Driver: Inbox, C-states: Disabled, Hyper Threading: Disabled, CPU Governor (through OS): Performance Mode , OS: CentOS 7.5.1804, Kernel: 4.14.74

Some results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

For copies of this document, documents that are referenced within, or other Intel literature, please contact your Intel representative.

Intel, Xeon and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2020, Intel Corporation. All rights reserved.

Contents

Intel® Solid State Drive D5-P4420 (U.2 Form Factor) **1**
Ordering Information **2**
Revision History **2**

1 Overview **8**
1.1 References 8
1.2 Terms and Acronyms 9

2 Product Specifications **11**
2.1 Capacity 11
2.2 Performance 11
2.3 Electrical Characteristics 11
2.4 Product Features and Availability 12
2.5 Endurance 12
2.6 Environmental Conditions 12
2.7 Product Regulatory Compliance 13
2.8 Reliability Specifications 15
2.9 Thermal Specifications 15
2.9.1 Thermal Throttling 15
2.9.2 Composite Temperature 16
2.9.3 Thermal Shutdown 17
2.9.4 Airflow Requirements 17
2.10 Power Loss Capacitor Test 17
2.11 Hot Plug Support 17
2.12 Out-of-Band Management (SMBus) 18
2.12.1 VPD Page Readout over SMBus 18

3 Mechanical Information **20**

4 Pin and Signal Descriptions **21**
4.1 Pin Signal Definitions 21

5 Supported Command Sets **23**
5.1 NVMe Admin Command Set 23
5.2 NVMe I/O Command Set 23
5.3 Object Missing 24
5.4 NVMe and Vendor Unique Log Page Support 24
5.5 SMART Attributes 29
5.6 SET Features Identifiers 31
5.7 Vendor Unique Opcodes 32
5.7.1 Get Intel Log (Opcode D2h) 32
5.8 Vendor Unique FIDs 32

6 NVMe Driver Support **35**

7 Other Compliance and Certifications **36**

Appendices **37**

A Performance and Endurance Metrics **37**

B Power Metrics....40

C IDENTIFY Data Structure....41

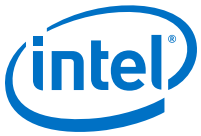
D Vital Data Structure (0x53)....57

E Out-of-Band Command Response Using SMBus (0x6A)....63

F Out-of-Band Command Response Using SMBus (0x6A Intel Specific)....65

G SCSI Command Translation....66

H PCIe* ID....67



Figures

Figure 1: Thermal Throttling Behavior....16
Figure 2: Intel® SSD D5-P4420 - Airflow Approach Curve....17
Figure 3: SMBus and VPD Connection....19
Figure 4: Intel® SSD D5-P4420 U.2 15mm Dimensions....20

Tables

Table 1: Standard Information Referenced in this Document ... 8
Table 2: Glossary of Terms and Acronyms ... 9
Table 3: Input Supply Rails ... 11
Table 4: Features and Availability ... 12
Table 5: Temperature, Shock, Vibration ... 12
Table 6: Product Regulatory Compliance Specifications ... 13
Table 7: Reliability Specifications ... 15
Table 8: Intel® SSD D5-P4420 Thermal Throttle Settings ... 16
Table 9: Out-of-Band Readout Address ... 18
Table 10: Pin Definition for U.2 Form Factor ... 21
Table 11: Log Page Directory (Log Identifier C0h) ... 24
Table 12: Read/Write Command Latency Log (Log Identifier C1h/C2h) ... 25
Table 13: Temperature[1] Statistics (Log Identifier C5h) ... 25
Table 14: Vendor Unique SMART Log (Log Identifier CAh) ... 25
Table 15: NVMe IO Queue Metrics Log Page (Log Identifier CBh) ... 28
Table 16: Drive Marketing Name Log (Log Identifier DDh) ... 28
Table 17: SMART Attributes (Log Identifier 02h) ... 29
Table 18: Get Log Page - Temperature Sensor Data Structure ... 31
Table 19: Get Intel Log - PRP Entry 1 ... 32
Table 20: Get Intel Log - PRP Entry 2 ... 32
Table 21: Get Intel Log - Command DWORD 10 ... 32
Table 22: Get Intel Log - Command DWORD 12 ... 32
Table 23: Get Intel Log - Command DWORD 13 ... 32
Table 24: Set Max LBA Setting - Command Dword 11 and Command Dword 12 ... 33
Table 25: Status Code - Set Max LBA Command Specific Status Values ... 33
Table 26: C6h - Set/Get Power (Typical) Governor Setting - Command Dword 11 ... 33
Table 27: Status Codes - Power Governor Setting Command Specific Status Value ... 33
Table 28: D5h - Reset Timed Workload Counters - Command Dword 11 ... 33
Table 29: E2h - Set/Get Enable Latency Tracking ... 33
Table 30: C8h - Get/Set SMB ASIC Address ... 33
Table 31: C9h - Set/Get Blink activity for LED ... 34
Table 32: D5-P4420 NVMe Driver Support ... 35
Table 33: Other Compliance and Certifications ... 36
Table 34: User Addressable Sectors (LBA) ... 37
Table 35: Random Read/Write Performance (IOPS)[1] ... 37
Table 36: Read/Write Consistency (%)[1] ... 37
Table 37: Sequential Read and Write Bandwidth (MB/s)[1] ... 38
Table 38: Latency[1] ... 38
Table 39: Quality of Service (QoS)[1] (µs) ... 38
Table 40: Endurance (DWPD) ... 38
Table 41: Endurance (PBW)[1] ... 39
Table 42: Power Consumption (W) ... 40
Table 43: Identify Controller ... 41
Table 44: Power State Descriptors Data Structure ... 49
Table 45: Vendor Specific Data Structure ... 50
Table 46: Identify Namespace Data Structure ... 51
Table 47: LBA Format Data Structure – 512B ... 56
Table 48: Possible DPC & DPS values ... 56
Table 49: Vital Product Data Structure (VPD Elements) ... 57
Table 50: Type/Length Byte Format ... 57
Table 51: Common Header ... 57
Table 52: Product Info Area (offset 8 bytes) ... 58
Table 53: NVMe MultiRecord Area ... 59

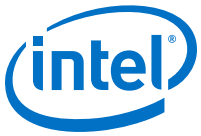


Table 54: NVMe PCIe Port MultiRecord Area....61
Table 55: System Management Data Structure (NVMe-MI Commands) 63
Table 56: Command Response 0x6A (Intel Specific Vendor Unique Commands)....65
Table 57: PCIe ID....67

1 Overview

This document describes the specifications and capabilities of the Intel® SSD D5-P4420(D5-P4420).

The D5-P4420 is PCIe* Gen3 SSD architected with Intel's first 3D QLC NAND technology and high performance controller interface – NVMe* (Non-Volatile Memory Express*) delivering leading performance, low latency and QoS (Quality of Service). Matching the cost and capacity optimization, Intel® SSD D5-P4420 is available in capacities of 15.36TB U.2 15mm form factor.

With PCIe Gen3 support and NVMe queuing interface, the D5-P4420 delivers excellent sequential read performance of up to 3.2GB/s and sequential write speeds of up to 1.6GB/s. The D5-P4420 delivers very high random read IOPS of 580K for 4KB operation and random write IOPS of 11K for 16KB operations. Taking advantage of the direct path from the storage to the CPU by means of NVMe, the D5-P4420 exhibits low latency for sequential access to the SSD.

The D5-P4420 includes these key features:

- Consistently High IOPS and throughput
- Sustained low latency
- Enhanced power-loss data protection
- Power loss protection capacitor self-test
- Out of band management
- Thermal throttling and monitoring
- SMART Health reporting
- End-to-end data protection

1.1 References

Table 1: Standard Information Referenced in this Document

Date	Title	Location
June 5, 2016	NVMe Revision 1.2.1	http://www.nvmexpress.org
Jan 2013	Enterprise SSD Form Factor Version 1.0a	http://www.ssdformfactor.org
Oct 2017	PCIe Base Specification Revision 4.0	http://pcisig.com
July 2012	Solid State Drive (SSD) Endurance Workloads (JESD219)	http://www.jedec.org/standards-documents/results/jesd219
Sept 2010	Solid State Drive (SSD) Requirements and Endurance Test Method (JESD218)	http://www.jedec.org/standards-documents/docs/jesd218/
Dec 2008	VCCI	http://www.vcci.jp/vcci_e/
June 2009	RoHS	http://qdms.intel.com/ Click *Search MDDS Database* and search for material description datasheet
1995 1996 1995	International Electrotechnical Commission EN 61000 4-2 (Electrostatic discharge immunity test) 4-3 (Radiated, radio-frequency, electromagnetic field immunity test)	http://www.iec.ch/

continued...

Date	Title	Location
1995 1997 1994	4-4 (Electrical fast transient/burst immunity test) 4-5 (Surge immunity test) 4-6 (Immunity to conducted disturbances, induced by radio-frequency fields) 4-11 (Voltage Variations, voltage dips, short interruptions and voltage variations immunity tests)	
1995	ENV 50204 (Radiated electromagnetic field from digital radio telephones)	http://www.dbicorporation.com/radimmun.htm/

1.2 Terms and Acronyms

Table 2: Glossary of Terms and Acronyms

Term	Definition
ATA	Advanced Technology Attachment
CRC	Cyclic Redundancy Check
DAS	Device Activity Signal
DMA	Direct Memory Access
ECC	Error Correction Code
EEPROM	Electrically Erasable Programmable Read Only Memory
GB	Gigabyte **Note:** The total usable capacity of the SSD may be less than the total physical capacity because a small portion of the capacity is used for 3D XPoint media management and maintenance purposes.
Gb	Gigabit
I/O	Input/ Output
IOPS	Input/ Output Operations Per Second
KB	Kilobyte
LBA	Logical Block Address
MB	Megabyte (1,000,000 bytes)
MI	Management Interface
MTBF	Mean Time Between Failures
NOP	No Operation
NVMe	Non-Volatile Memory Express
PB	Petabyte
PCB	Printed Circuit Board
RDT	Reliability Demonstration Test
RMS	Root Mean Square
SMART	Self-Monitoring, Analysis and Reporting Technology. This is an open standard for developing hard drives and software systems that automatically monitors the health of a drive and reports potential problems.
SRIS	Separate Refclk Independent Spread
SSD	Solid State Drive

continued...

Term	Definition
TB	Terabyte
TBD	To Be Determined
TYP	Typical
UBER	Uncorrectable Bit Error Rate
VPD	Vital Product Data

§

2 Product Specifications

2.1 Capacity

Total user addressable sectors (LBAs) per drive capacity information is provided in Appendix A: Performance and Endurance Metrics.

2.2 Performance

Performance specifications are highly dependent on the product operating within the operating requirements as listed in the Environmental Conditions section. Additional information is provided in Appendix A: Performance and Endurance Metrics.

2.3 Electrical Characteristics

Table 3: Input Supply Rails

Electrical Characteristics	12V Host	3.3V Aux[4]
Tolerance	+10% -20%	±9%
Max Startup Current[1] (first 2s operating) Inrush Current (Typical Peak)[1]	2A	-
Max Average Current[2]	2.45A (2.1A)	25mA active
Min Off-Time	500ms	500ms
Rising Slew Rate	250V/s – 420kV/s	250V/s – 420kV/s
Falling Slew Rate[3]	2.4 V/s - 250kV/s	0.66V/s - 33kV/s
Shutdown Undershoot	0V (No undershoot allowed)	0V (No undershoot allowed)
Noise 10Hz – 100kHz	1000mV	300mV
Noise 100kHz – 20MHz	50mV	50mV

Notes:

1. Measured during initial power supply application. Typically this will be seen within 2 seconds of initial power up. Inrush specified for 12V and 3.3V supply, not the 3.3Vaux.
2. PCB design shall support an input supply voltage of 12V + 10, -20%. If the supply voltage is held within the PCIe CEM spec [+/-8%], average current limit of [2.1A] shall not be exceeded.
3. Fall time needs to be equal or better than minimum in order to guarantee full functionality of enhanced power loss management.
4. 3.3Vaux is optional, not needed for power up or functionality. 3.3Vaux is needed for accessing VPD page by means of SMBUS.

2.4 Product Features and Availability

Table 4: Features and Availability

Features	Availability
Common Clock (RefClk)	Production Release
Crypto Erase, User Data Erase, FormatNVM	Production Release
SMART	Production Release
Firmware commit with NVMe Controller Reset	Production Release
Firmware Image Download	Production Release
U.2 Hot Plug	Production Release
Thermal shutdown	Production Release
Composite Temperature	Production Release
Dataset Management (Deallocate)	Production Release
Multiple Power Modes (Vendor Unique, refer to Section 5.7)	Production Release
Multiple Namespace Support Namespace Attach Namespace Management	Future Maintenance Release
Crypto Erase per Namespace Physical Erase per Namespace	Future Maintenance Release
NVMe-MI Full (SMBus)	Production Release
Firmware activate without RESET	Future Maintenance Release
SRIS	EDSFF form factor
OPAL 2.0	Future Maintenance Release
Write Zeroes	Future Maintenance Release
Multiple Firmware Slots	Future Maintenance Release

2.5 Endurance

Additional information is provided in Appendix A: Performance and Endurance Metrics.

2.6 Environmental Conditions

Power information is provided in Appendix B: Power Metrics.

Table 5: Temperature, Shock, Vibration

Specification		
Temperature	Temperature Operating[1] Non-operating[2]	0° C to 70° C -55° C to 95° C
	Temperature Gradient[3] Operating Non-operating	20° C/hr (Typical) 30° C/hr (Typical)
Humidity	Humidity Operating Non-operating	5 – 90% 5 – 95%

continued...

Specification		
Shock and Vibration	Shock[4] Operating Non-operating	1,000 G (Max) at 0.5 msec 1,000 G (Max) at 0.5 msec
	Vibration[5] Operating Non-operating	2.17 G_{RMS} (5 - 700 Hz) Max 3.13 G_{RMS} (5 - 800 Hz) Max
Altitude	Operating Non-operating	-1,000 to 10,000ft. -1,000 to 40,000ft.

Notes:

1. Operating temperature implies ambient air temperature under defined airflow. Operation temperature refers to SMART temperature.
2. Please contact your Intel representative for details on the non-operating temperature range.
3. Temperature gradient measured without condensation.
4. Shock specifications assume the SSD is mounted securely with the input vibration applied to the drive-mounting screws. Stimulus may be applied in the X, Y or Z axis. Shock specification is measured using Root Mean Squared (RMS) value.
5. Vibration specifications assume the SSD is mounted securely with the input vibration applied to the drive-mounting screws. Stimulus may be applied in the X, Y or Z axis. Vibration specification is measured using RMS value.

2.7 Product Regulatory Compliance

Intel® SSD D5-P4420 drive meet or exceeds the regulatory or certification requirements in the following table. The following table applies to the production unit only.

Table 6: Product Regulatory Compliance Specifications

Certification	Description
CE Compliant	European Economic Area (EEA): Compliance with the essential requirements of EC Council Directives Low Voltage Directive (LVD) 2014/35/EU and EMC Directive 2014/30/EU. Compliance with Directive 2011/65/EU of the European Parliament and of the Council of 8 June 2011 on the restriction of the use of certain hazardous substances in electrical and electronic equipment
EU WEEE	Compliance with Directive 2012/19/EU of the European Parliament and of the Council of 4 July 2012 on waste electrical and electronic equipment (WEEE)
UL Recognized	Certified Underwriters Laboratories, Inc. Bi-National Component Recognition; UL 60950-1, 2nd Edition, 2014-10-14 [Information Technology Equipment - Safety - Part 1: General Requirements]. CSA C22.2 No. 60950-1-07, 2nd Edition, 2014-010 (Information Technology Equipment - Safety - Part 1: General Requirements) These products have been Complimentary Recognized to UL/CSA 62368-1, 2nd Edition [Audio/video, information and communication technology equipment - Part 1: Safety requirements] **"CAUTION: Hot Surface, Do not touch the SSD surface "**

continued...

Certification	Description
Australia / New Zealand: RCM	Compliance with the Australia/New Zealand Standard(s) AS/NZ CISPR 32:2015 and AS/NZ CISPR 22:2009 +A1:2010, In compliance with the Radiocommunications Act 1992 as part of the ACMA's Electromagnetic Compatibility (EMC) Regulatory Arrangement and RSM Radiocommunications (EMC Standards) Notice 2015.
Taiwan BSMI	Compliance to the Taiwan EMC standard CNS 13438: Information technology equipment - Radio disturbance Characteristics - limits and methods of measurement, as amended on June 1, 2006, is harmonized with CISPR 22: 2005.04. Compliance to the Taiwan CNS 15663 [Guidance to reduction of the restricted chemical substances in electrical and electronic equipment (EEE)].
Korea KCC	Compliance with paragraph 1 of Article 11 of the Electromagnetic Compatibility Control Regulation and meets the Electromagnetic Compatibility (EMC) Framework requirements of the Radio Research Laboratory (RRL) Ministry of Information and Communication Republic of Korea.
Morocco Maghreb	Compliant with Decree # 2574-14 (EMC) on electromagnetic compatibility
Canada ICES **CAN ICES-3 (A)/ NMB-3(A)**	Compliance with Innovation, Science and Economic Development Canada standard ICES-003.
Japan VCCI	Voluntary Control Council for Interface to cope with disturbance problems caused by personal computers or facsimile.
China EFUP	China Environmentally Friendly Use Period (EFUP) symbol. Compliance with GB/T 26572: Requirements on concentration limits for certain restricted substances in electrical and electronic products.
Ukraine RoHS	Compliant with TECHNICAL REGULATION on the restriction of the use of certain hazardous substances in electrical and electronic equipment APPROVED by Resolution of the Cabinet of Ministers of Ukraine of 10 March 2017 No. 139.
Low Halogen	Applies only to brominated and chlorinated flame retardants (BFRs/CFRs) and PVC in the final product. Intel components as well as purchased components on the finished assembly meet JS-709 requirements, and the PCB/ substrate meet IEC 61249-2-21 requirements. The replacement of halogenated flame retardants and/or PVC may not be better for the environment.

2.8 Reliability Specifications

Intel® SSD D5-P4420 meets or exceeds SSD endurance and data retention requirements as specified in the JESD219 standard. Reliability specifications are listed in the table below.

Table 7: Reliability Specifications

Parameter	Value
Uncorrectable Bit Error Rate (UBER) • Uncorrectable bit error will not exceed one sector in the specified number of bits read. In the unlikely event of an non-recoverable read error, the SSD will report it as a read failure to the host; the sector in error is considered corrupt and is not returned to the host.	< 1 sector per 10^{17} bits read
Mean Time Between Failures (MTBF) • Mean Time Between Failures is estimated based on Telcordia* methology and demonstrated through Reliability Demonstration Test (RDT).	2 million hours
Data Retention • The time period for retaining data in the NAND at maximum rated endurance	3 months power-off retention once SDD reaches rated write endurance at 40° C

Note: Refer to JESD219 standard table 1 for UBER, FFR, and other Enterprise SSD requirements.

2.9 Thermal Specifications

2.9.1 Thermal Throttling

Intel® SSD D5-P4420 will provide performance throttling during high temperature scenarios to mitigate thermal challenges. Thermal throttling works by moving the power ceiling in N steps as shown in Figure 1.

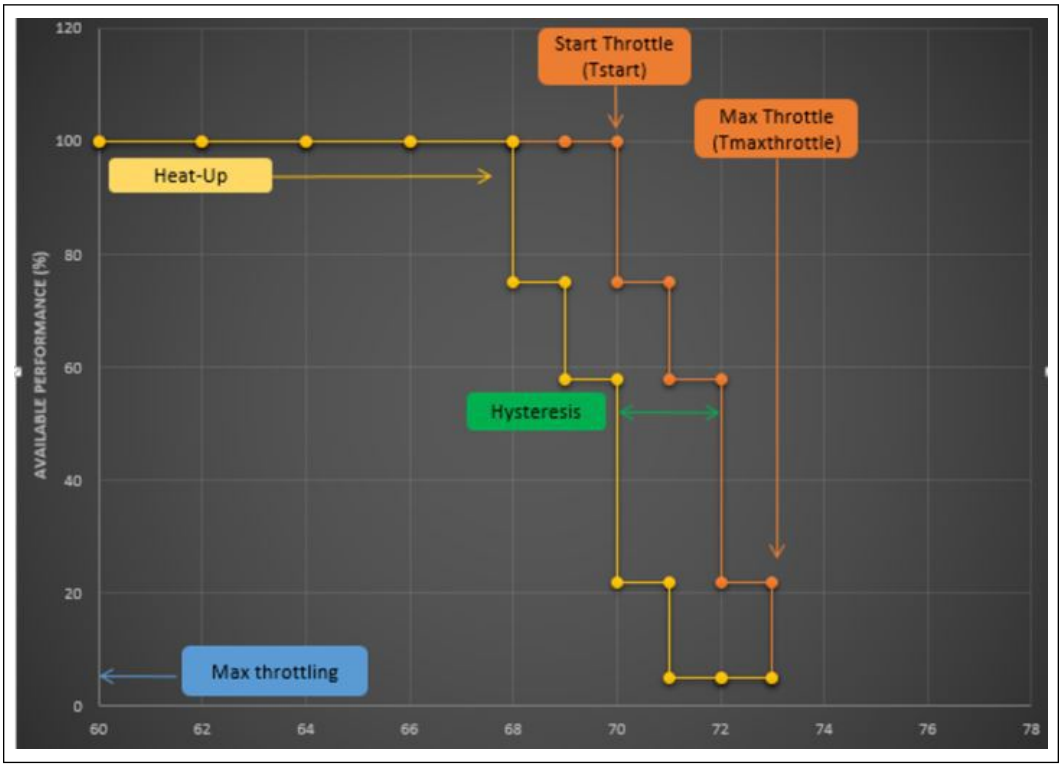
Key definitions for important parameters are:

-Tstart - throttling starts

-Tmaxthrottle - max throttling is applied

-Hysteresis - 2° C comes into play during the cooling phase to prevent rapid oscillations between the throttle states. The temperature must hit the hysteresis level to return the power ceiling back to previous levels.

Figure 1: Thermal Throttling Behavior



Shown above is the Thermal throttling behavior for Intel® SSD D5-P4420, T_{start} set to 70° C Composite Temperature, $T_{maxthrottle}$ is set to 73° C, Hysteresis is set to 2° C and N is set to 4

Table 8: Intel® SSD D5-P4420 Thermal Throttle Settings

Form Factor	Capacity (TB)	Throttle Start	Throttle Max	Throttle Steps
U.2 15mm	7.68TB	70° C	73° C	4

2.9.2 Composite Temperature

The Intel® SSD D5-P4420 follows the composite temperature scheme as part of NVMe Health log (SMART attribute log page identifier 02h, byte 1&2) which takes into account, ASIC, Board and NAND temperature sensors and provide an aggregate value. This value represents all components with temperature monitoring capability, resolving all components to a single critical temperature upon which throttling is based. Thermal throttle gets engaged when the composite temperature reaches 70° C (as reported by SMART). For more information on sensor reading see SMART attributes section.

The focus of this thermal management system is to utilize temperature sensors now integrated onto the media die and other parts of the drive PCB. This allows monitoring of key component temperature in the varied environment that leads to more precise thermal management.

In addition, drive will provide out of band access to temperature via SMBUS. The SMBus slave address to read SMART data structure is the same address we use for MCTP, and defaults to 0x6Ah.

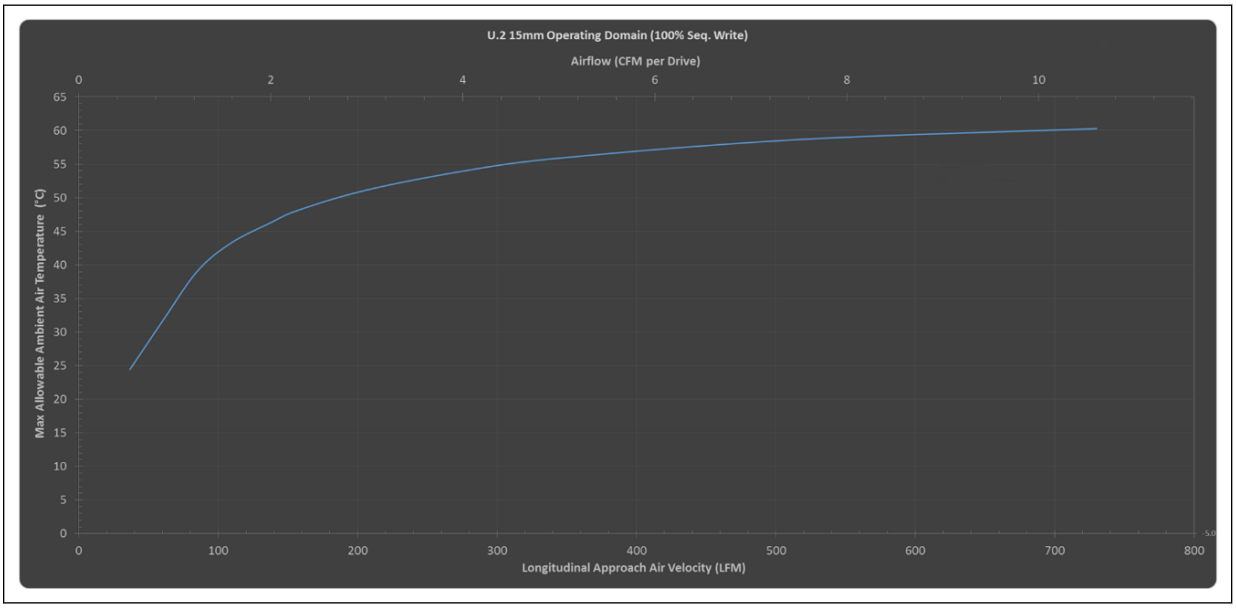
The NVMe 1.2.1 spec recommends a Warning Composite Temperature Threshold (WCTEMP) of 70° C.

2.9.3 Thermal Shutdown

The Intel® SSD D5-P4420 supports thermal shutdown at a SMART temperature of 80° C. If thermal throttling fails to maintain a suitable temperature, the drive will trigger a PLI event then completely shut down.

2.9.4 Airflow Requirements

Figure 2: Intel® SSD D5-P4420 - Airflow Approach Curve



Shown above is the CFM per drive and Approach LFM vs. Ambient Air Temperature for Intel® SSD D5-P4420 U.2 15mm 7.68TB

Above curve represents the CFM per drive and approach LFM required to prevent drive from throttling at a given ambient air temperature. The curve are based on longitudinal flow orientation with U.2 connector downstream. Measurement data collected to plot the curve is in a typical data center drive bay configuration with 18mm pitch which results in 3mm spacing between drives . The curve is subject to change with change in hardware, system configuration and drive power

2.10 Power Loss Capacitor Test

The Intel® SSD D5-P4420 supports testing of the power loss capacitor, which can be monitored using SMART attribute critical warning in log page identifier 02h, byte 0, and bit 4.

2.11 Hot Plug Support

The U.2 form factor SSDs support orderly hot insertion and removal and surprise hot insertion by means of presence detect and link-up detect in capable platforms and OSs. On surprise hot removal during IOs, the drive will ensure the integrity of already committed data on the media and commit acknowledged writes to the media.

2.12 Out-of-Band Management (SMBus)

D5-P4420 provides out-of-band management by means of an SMBus interface on two addresses; 0x53/0xA6 (7bit/8bit) provides a Vital Product Data (VPD) page; and 0x6A/0xD4 (7bit/8bit) provides an NVMe-MI 1.0 interface.

The VPD page may be read when 3.3Vaux is applied to the drive. It does not require the drive to have 12V applied nor a properly configured PCIe link. More details on the structure of the VPD page at address 0x53/0xA6 (7bit/8bit) can be found in the Vital Data Structure appendix.

Table 9: Out-of-Band Readout Address

Description	Address	
	7-bit Address	8-bit Address
VPD Page	0x53	0xA6
NVMe MI Basic	0x6A	0xD4

D5-P4420 provides additional drive information via address 0x6A/0xD4 (7bit/8bit) as outlined in the NVMe Basic Management Command (see here). This interface requires the drive to have power supplied, a valid PCIe link, an initialized NVMe configuration space, and the controller must be enabled. Available commands on this interface are detailed in Appendices Out-of-Band Command Response Using SMBus (0x6A) on page 63 and Out-of-Band Command Response Using SMBus (0x6A Intel Specific) on page 65, which have details on the Out-of-Band Management data structure.

Notes:

1. In certain tools the address for the VPD and temperature sensor will appear as 0xA6 and 0xD4 respectively, due to bit shift.
2. BMC should not access the SMBUS address within 100msec of device power up, it may experience some glitch on the bus. Some of the drive information (NVMe MI Appendix A per NVMe MI Specification) drive functional status, temperature read out, etc. might have stale value right after power ON and host should wait until drive TTR recommendation. To see the Time to Ready (TTR) refer to the Latency table in Appendix A.

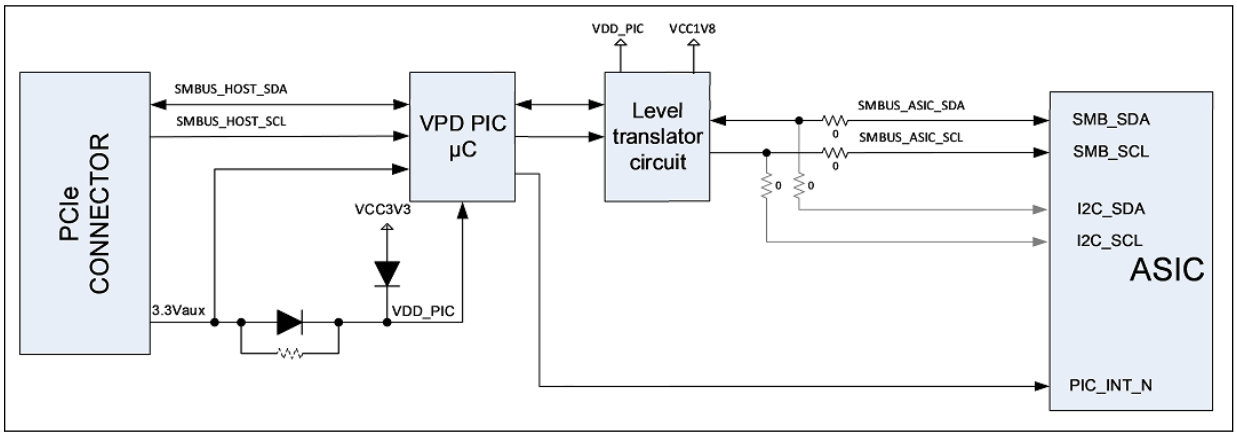
2.12.1 VPD Page Readout over SMBus

D5-P4420 should support simple Reads to Vital Product Data (VPD). Please refer to Vital Data Structure (0x53) on page 57 for details on VPD Data Structure. VPD contains:

- Basic inventory information such as type and size of Enterprise PCIe SSD, manufacturer, date, revision, and GUID
- Power management data such as power level and power modes
- Vendor specific data

VPD is stored in an SMBus device with a slave address of 0xA6 (i.e., slave address bits 7-1 correspond to 1010_011). The VPD page can be read via SMBus through address 0x 53 (7- bit address with bit 8 set to 1) or 0xA6 (8-bit address).

Figure 3: SMBus and VPD Connection

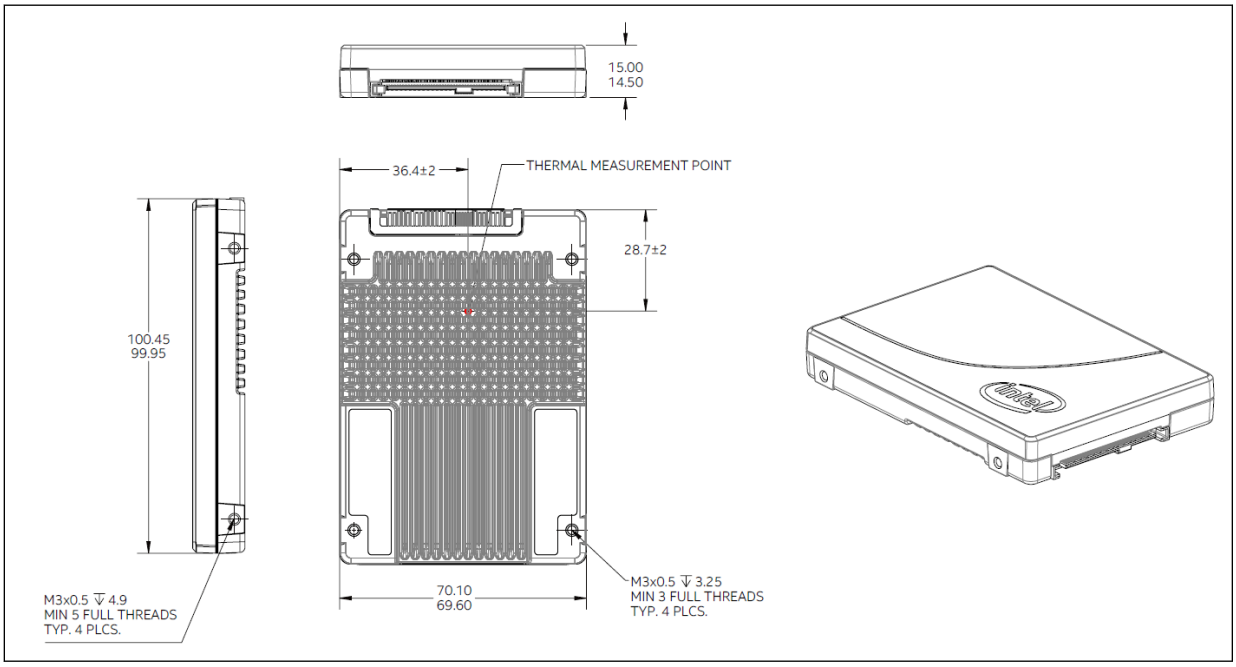


§

3 Mechanical Information

The following figures show the physical package information for the Intel® SSD D5-P4420 in the U.2 15mm form factor. All dimensions are in millimeters.

Figure 4: Intel® SSD D5-P4420 U.2 15mm Dimensions



X - Length	Y - Width	Z - Height
100.45 Max	69.85 +/- 0.25	15.0 +0/-0.5

Note: Length does not include 0.3 connector protrusion.

§

4 Pin and Signal Descriptions

4.1 Pin Signal Definitions

Table 10: Pin Definition for U.2 Form Factor

Pin	Name	Description	Pin	Name	Description
S1	GND	Ground	E7	REFCLK0+	Reference clock port 0
S2		Not used (SATA/SAS)	E8	REFCLK0-	Reference clock port 0
S3		Not used (SATA/SAS)	E9	GND	Ground
S4	GND	Ground	E10	PETp0	Transmitter differential pair, Lane 0
S5		Not used (SATA/SAS)	E11	PETn0	Transmitter differential pair, Lane 0
S6		Not used (SATA/SAS)	E12	GND	Ground
S7	GND	Ground	E13	PERn0	Receiver differential pair, Lane 0
E1	REFCLK1+	Reference clock port 1	E14	PERp0	Receiver differential pair, Lane 0
E2	REFCLK1-	Reference clock port 1	E15	GND	Ground
E3	3.3Vaux	3.3V auxiliary power	E16	RSVD	Reserved
E4	PERST1#	Fundamental reset port 1	S8	GND	Ground
E5	PERST0#	Fundamental reset port 0	S9		Not used (SATAe/SAS)
E6	RSVD	Reserved	S10		Not used (SATAe/SAS)
P1		Not used (SATAe/SAS)	S11	GND	Ground
P2		Not used (SATAe/SAS)	S12		Not used (SATAe/SAS)
P3		Not used (SATAe)	S13		Not used (SATAe/SAS)
P4	IfDet_N1	Interface detect (drive type)	S14	GND	Ground
P5	GND	Ground	S15	RSVD	Reserved
P6	GND	Ground	S16	GND	Ground
P7		Not used (SATA/SAS)	S17	PETp1	Transmitter differential pair, Lane 1
P8		Not used (SATA/SAS)	S18	PETn1	Transmitter differential pair, Lane 1
P9		Not used (SATA/SAS)	S19	GND	Ground
P10	PRSNT_N	Presence detect (also used for drive type)	S20	PERn1	Receiver differential pair, Lane 1
P11	Activity_N[1]	Activity signal from the drive	S21	PERp1	Receiver differential pair, Lane 1
P12	Hot-Plug	Ground	S22	GND	Ground
P13	+12V_pre	12V power	S23	PETp2	Transmitter differential pair, Lane 2
P14	+12V	12V power	S24	PETn2	Transmitter differential pair, Lane 2
P15	+12V	12V power	S25	GND	Ground

continued...

Pin	Name	Description	Pin	Name	Description
			S26	PERn2	Receiver differential pair, Lane 2
			S27	PERp2	Receiver differential pair, Lane 2
			S28	GND	Ground
			E17	PETp3	Transmitter differential pair, Lane 3
			E18	PETn3	Transmitter differential pair, Lane 3
			E19	GND	Ground
			E20	PERn3	Receiver differential pair, Lane 3
			E21	PERp3	Receiver differential pair, Lane 3
			E22	GND	Ground
			E23	SMCLK[2]	SMBUS clock
			E24	SMDAT[2]	SMBUS data
			E25	DualPortEn_N	Dual port enable

§

5 Supported Command Sets

Intel® SSD D5-P4420 supports all mandatory Admin and I/O commands defined in NVMe 1.2.1 specification.

5.1 NVMe Admin Command Set

The Intel® SSD D5-P4420 supports all mandatory NVMe commands, which are:

- Delete I/O Submission Queue
- Delete I/O Completion Queue
- Create I/O Submission Queue
- Create I/O Completion Queue
- Get Log Page
- Identify
- Abort
- SET Features
- GET Features
- Asynchronous Event Request

D5-P4420 also supports the following optional I/O commands defined in NVMe specification revision 1.2.1

- Firmware Activate
- Firmware Image Download
- Format NVM
- Namespace Management[1]
- Namespace Attachment[1]
- Security Send[2]
- Security Receive [2]

Notes: See Appendix IDENTIFY Data Structure for details on commands and capabilities.

1. On drives with MR5 firmware
2. Only on OPAL SKUs

5.2 NVMe I/O Command Set

D5-P4420 supports all of the mandatory NVMe I/O command set defined in NVMe 1.2.1 specification, which includes:

- Flush
- Write
- Read

Additionally, the following optional commands are supported:

- Write Uncorrectable

- Dataset Management (Deallocate only)

Note: See Appendix SCSI Command Translation on page 66 for details on SCSI supported commands and capabilities.

5.3 Object Missing

This object is not available in the repository.

5.4 NVMe and Vendor Unique Log Page Support

D5-P4420 supports the following mandatory log pages defined in NVMe 1.2.1 specification:

- Error Information (Log Identifier 01h)
- SMART/ Health Information (Log Identifier 02h)
- Firmware Slot Information (Log Identifier 03h)
- Command Effects Log (Log Identifier 05h)

Note: See NVMe 1.2.1 version of the NVMe specification for the log page content.

Additionally, D5-P4420 will support the following Vendor Unique (C0h-FFh) log pages:

- Log Page Directory (Log Identifier C0h)
- Read Command Latency Statistics Log Page (Log Identifier C1h)
- Write Command Latency Statistics Log Page (Log Identifier C2h)
- Temperature Statistics (Log Identifier C5h)
- Vendor Unique SMART Log (Log Identifier CAh)
- Vendor Unique NVMe IO Queue Metrics Log Page Log Identifier CBh)
- Marketing Description Log (Log Identifier DDh)

Table 11: Log Page Directory (Log Identifier C0h)

Intel® SSD D5-P4420		
Byte	**# of Bytes**	**Log Page Content**
0-1	2	Log version
386	1	Number of 512B Log Pages at Log Address C1h
388	1	Number of 512B Log Pages at Log Address C2h
394	1	Number of 512B Log Pages at Log Address C5h
404	1	Number of 512B Log Pages at Log Address CAh
406	1	NVMe IO Queue Metrics Log Page at Log Address CBh
442	1	Number of 512B Log Pages at Log Address DDh

Log page C0h will hold the above mentioned entities, describing vendor unique Log pages that the device supports. The structure below follows ATA structure, to provide consistency. The directory shows 5 VU Log pages supported. Each Log page is 512B long. Details of individual Log pages can be found under their respective definitions in the tables below.

Table 12: Read/Write Command Latency Log (Log Identifier C1h/C2h)

Intel® SSD D5-P4420		
Byte	**# of Bytes**	**Log Page Content**
0-1	2	Major Version
2-3	2	Minor Version
4-4867	4864	1216 buckets, each bucket size is 4 bytes Full range is range 0 - 2s and is logarithmic scaling ~10% max error per bucket

These log pages will show zero contents until latency tracker is enabled using the Set Features command E2h (Set/Get Enable Latency Tracking). Enabling latency tracker adds a performance penalty and must be disabled upon completing the debug.

Table 13: Temperature[1] Statistics (Log Identifier C5h)

Intel® SSD D5-P4420		
Byte	**# of Bytes**	**Log Page Content**
0-7	8	Current Composite Temperature in Celsius
8-15	8	SSD Overtemp Shutdown flag for last power on. Clears after subsequent reset.
16-23	8	SSD Overtemp Shutdown flag for lifetime. Sticky/persistent once set.
24-31	8	Highest (Lifetime) Composite Temperature in Celsius
32-39	8	Lowest (Lifetime) Composite Temperature in Celsius
40-79	40	Reserved
80-87	8	Max Warning Normalized Threshold in Celsius (TX)
88-95	8	Reserved
96-103	8	Specified Minimum Operating Temp in Celsius (TM)
104-111	8	Estimated Offset in Celsius (TC)
111-511	400	Reserved

Note: 1. All temperature values indicate internal composite temperature values. The log page will read 00h for reserved bytes.

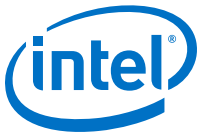
Table 14: Vendor Unique SMART Log (Log Identifier CAh)

Byte Offset	Attribute	Description	Updated	Saved to NAND
00h	**AB (Program Fail Count)**	**Raw value:** shows total count of program fails. **Normalized value:** beginning at 100, shows the percent remaining of allowable program fails.	Upon Event	Upon Event
01h	Reserved			
03h	Normalized Value			
04h	Reserved			
05h	Current Raw Value			
0Ch	**AC (Erase Fail Count)**	**Raw value:** shows total count of erase fails.	Upon Event	Upon Event

continued...

Byte Offset	Attribute	Description	Updated	Saved to NAND
0Dh	Reserved	**Normalized value:** beginning at 100, shows the percent remaining of allowable erase fails.		
0Fh	Normalized Value			
10h	Reserved			
11h	Current Raw Value			
18h	**AD (Wear Leveling Count)**	**Raw value:** Min, max and average values of NAND erase cycles for all blocks. Bytes 1-0: Min. erase cycle Bytes 3-2: Max. erase cycle Bytes 5-4: Avg. erase cycles **Normalized value:** decrements from 100 to 0.	Upon Event	Upon Event
19h	Reserved			
1Bh	Normalized Value			
1Ch	Reserved			
1Dh	Current Raw Value			
24h	**B8 (End to End Error Detection Count)**	**Raw value:** reports number of End-to-End detected and corrected errors by hardware. **Normalized value:** always 100.	Upon Event	Upon Event
25h	Reserved			
27h	Normalized Value			
28h	Reserved			
29h	Current Raw Value			
30h	**C7 (CRC Error Count)**	**Raw value:** total number of PCIe Interface CRC errors encountered, as specified in PCIe Link Performance Counter Parameter for "Bad TLP". Count is preserved across power cycles. **Normalized value:** always 100.	Upon Event	Upon Event
31h	Reserved			
33h	Normalized Value			
34h	Reserved			
35h	Current Raw Value			
3Ch	**E2 (Timed Workload, Media Wear)**	**Raw value:** measures the wear seen by the SSD (since reset of the workload timer, attribute E4h), as a percentage of the maximum rated cycles. Divide the raw value by 1024 to derive the percentage with 3 decimal points. Example: if the raw value is 4450, the percentage is 4450/1024 = 4.345%. Normalized value: always 100.	After 60 Minutes operation	Every 60 seconds and upon shutdown notification
3Dh	Reserved			
3Fh	Normalized Value			
40h	Reserved			
41h	Current Raw Value			
48h	**E3 (Timed Workload, Host Reads %)**	**Raw value:** shows the percentage of I/O operations that are read operations (since reset of the workload timer, attribute E4h). Reported as integer percentage from 0 to 100. **Normalized value:** always 100.	After 60 Seconds operation	Every 60 seconds and upon shutdown notification
49h	Reserved			
4Bh	Normalized Value			
4Ch	Reserved			
4Dh	Current Raw Value			

continued...



Byte Offset	Attribute	Description	Updated	Saved to NAND
54h	**E4 (Timed Workload, Timer)**	**Raw value:** measures the elapsed time (number of minutes since starting this workload timer). Example: if the raw value is 500, the timer has been running for 500 minutes. **Normalized value:** always 100.	After 60 Seconds operation	Every 60 seconds and upon shutdown notification
55h	Reserved			
57h	Normalized Value			
58h	Reserved			
59h	Current Raw Value			
60h	**EA (Thermal Throttle Status)**	**Raw value:** reports Percent Throttle Status and Count of events Byte 0: Throttle status reported as integer percentage. Bytes 1-4: Throttling event count. Number of times thermal throttle has activated. Preserved over power cycles. Byte 5: Reserved. **Normalized value:** always 100.	When Requested	Upon Event
61h	Reserved			
63h	Normalized Value			
64h	Reserved			
65h	Current Raw Value			
6Ch	**F0 (Retry Buffer Overflow Counter)**	**Raw Value:** Counter to indicate the number of times Retry Buffer has overflown **Normalized Value:** is always 100	Upon Event	Upon Event
6Dh	Reserved			
6Fh	Normalized Value			
70h	Reserved			
71h	Current Raw Value			
78h	**F3 (PLL Lock Loss Count)**	**Raw Value:** Counter to indicate the number of times PCIe Refclock PLL has unlocked **Normalized Value:** is always 100	Upon Event	Upon Event
79h	Reserved			
7Bh	Normalized Value			
7Ch	Reserved			
7Dh	Current Raw Value			
84h	**F4 (NAND Bytes Written)**	**Current Value:** NAND sectors written divided by 65536 (1 count = 32 MiB) **Normalized value:** always 100	When Requested	When requested but no more frequently than once per minute
85h	Reserved			
87h	Normalized Value			
88h	Reserved			
89h	Current Value			
90h	**F5 (Host Bytes Written)**	**Current Value:** Host sectors written divided by 65536 (1 count = 32 MiB) **Normalized value:** always 100	When Requested	When Requested
91h	Reserved			
93h	Normalized Value			
94h	Reserved			
95h	Current Value			

continued...

Byte Offset	Attribute	Description	Updated	Saved to NAND
9Ch	**F6 (System Area Life Remaining)**	**Current value** is normalized, representing the amount of system area writes that have been utilized. A value of 0x64 / 100d reflects that the user's allocation of system area writes has been exhausted.	When Requested	When Requested
9Dh	Reserved			
9Fh	Normalized Value			
A0h	Reserved			
A1h	Current Value			

Note: It is recommended for the host to wait at least 90seconds after power ON to retrieve the most current counter values from SMART Log Attributes (Log CAh)

Table 15: NVMe IO Queue Metrics Log Page (Log Identifier CBh)

Offset	# of Bytes	Log Page Content
0-1	2	Log Version (current is 1)
2-3	2	IOSQ Count
4-5	2	IOCQ Count
6-389	384	IOSQ Structs (12B each @ 32 queues)
390-709	320	IOCQ Structs (10B each @ 32 queues)
710-1023	314	Reserved padding to reach 1024B
IOSQ Structure Definition		
0-1	2	IOSQ ID
2-3	2	Associated IOCQ ID
4-5	2	Head Pointer
6-7	2	Tail Pointer
8-9	2	Outstanding Commands
10-11	2	Queue Size (Max Depth)
IOCQ Structure Definition		
0-1	2	IOCQ ID
2-3	2	Head Pointer

Table 16: Drive Marketing Name Log (Log Identifier DDh)

Byte	# of Bytes	Log Page Content
0	8	Intel®
8	1	Space
9	3	SSD
12	1	Space
13	2	DC
15	1	Space

continued...

Byte	# of Bytes	Log Page Content
16	5	Product Name
21	3	Space (three)
24	6	Series
30-511	482	Reserved (0x0)

5.5 SMART Attributes

The following tables list the SMART attributes supported by the D5-P4420 in accordance with NVMe 1.2.1 specification.

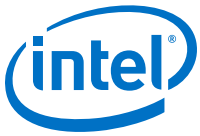
Table 17: SMART Attributes (Log Identifier 02h)

Byte	# of Bytes	Attribute	Description
0	1	Critical Warning	These bits if set, flag various warning sources. Bit 0: Available Spare is below Threshold Bit 1: Temperature has exceeded Threshold Bit 2: Reliability is degraded due to excessive media or internal errors Bit 3: Media is placed in Read-Only Mode Bit 4: Volatile Memory Backup System has failed (e.g., enhanced power loss capacitor test failure) Bits 5-7: Reserved Any of the critical warning can be tied to asynchronous event notification.
1	2	Temperature	Overall Device current temperature in Kelvin. For M.2 and U.2 7mm form factors, it is the composite temperature.
3	1	Available Spare	Starts from 100 and decrements.
4	1	Available Spare Threshold	Contains a normalized percentage (0 to 100%) of the remaining spare capacity available Threshold is set to 10%.
5	1	Percentage Used Estimate (Value allowed to exceed 100%)	A value of 100 indicates that the estimated endurance of the device has been consumed, but may not indicate a device failure. The value is allowed to exceed 100. Percentages greater than 254 shall be represented as 255. This value shall be updated once per power-on hour (when the controller is not in a sleep state).
32	16	Data Units Read (in LBAs)	Contains the number of 512 byte data units the host has read from the controller; this value does not include metadata. This value is reported in thousands (i.e., a value of 1 corresponds to 1000 units of 512 bytes read) and is rounded up. When the LBA size is a value other than 512 bytes, the controller shall convert the amount of data read to 512 byte units.
48	16	Data Units Write (in LBAs)	Contains the number of 512 byte data units the host has written to the controller; this value does not include metadata. This value is

continued...

Byte	# of Bytes	Attribute	Description
			reported in thousands (i.e., a value of 1 corresponds to 1000 units of 512 bytes written) and is rounded up. When the LBA size is a value other than 512 bytes, the controller shall convert the amount of data written to 512 byte units. For the NVM command set, logical blocks written as part of Write operations shall be included in this value. Write Uncorrectable commands shall not impact this value.
64	16	Host Read Commands	Contains the number of read commands issued to the controller.
80	16	Host Write Commands	Contains the number of write commands issued to the controller.
96	16	Controller Busy Time (in minutes)	Contains the amount of time the controller is busy with I/O commands. The controller is busy when there is a command outstanding to an I/O Queue (specifically, a command was issued by way of an I/O Submission Queue Tail doorbell write and the corresponding completion queue entry has not been posted yet to the associated I/O Completion Queue). This value is reported in minutes.
112	16	Power Cycles	Contains the number of power cycles.
128	16	Power On Hours	Contains the number of power-on hours. This does not include time that the controller was powered and in a low power state condition.
144	16	Unsafe shutdowns	Contains the number of unsafe shutdowns. This count is incremented when a shutdown notification (CC.SHN) is not received prior to loss of power.
160	16	Media Errors	Contains the number of occurrences where the controller detected an unrecovered data integrity error. Errors such as uncorrectable ECC, CRC checksum failure, or LBA tag mismatch are included in this field.
176	16	Number of Error Information Log Entries	Contains the number of Error Information log entries over the life of the controller.
195:192	4	Warning Composite Temperature Time	Contains the amount of time in minutes that the controller is operational and the Composite Temperature is greater than or equal to the Warning Composite Temperature Threshold (WCTEMP) field and less than the Critical Composite Temperature Threshold (CCTEMP) field in the Identify Controller data structure in Appendix C. If the value of the WCTEMP or CCTEMP field is 0h, then this field is always cleared to 0h regardless of the Composite Temperature value.
199:196	4	Critical Composite Temperature Time	Contains the amount of time in minutes that the controller is operational and the Composite Temperature is greater the Critical Composite Temperature Threshold (CCTEMP) field in the Identify Controller data structure in Figure 111 of the *NVM Express Base Specification*.

continued...



Byte	# of Bytes	Attribute	Description
			If the value of the CCTEMP field is 0h, then this field is always cleared to 0h regardless of the Composite Temperature value
201:200	16	Temperature Sensor 1	Not currently implemented in D5-P4420
203:202	16	Temperature Sensor 2	Not currently implemented in D5-P4420
205:204	16	Temperature Sensor 3	Not currently implemented in D5-P4420
207:206	16	Temperature Sensor 4	Not currently implemented in D5-P4420
209:208	16	Temperature Sensor 5	Not currently implemented in D5-P4420
211:210	16	Temperature Sensor 6	Not currently implemented in D5-P4420
213:212	16	Temperature Sensor 7	Not currently implemented in D5-P4420
215:214	16	Temperature Sensor 8	Not currently implemented in D5-P4420
511:216	N/A	Reserved	

Table 18: Get Log Page - Temperature Sensor Data Structure

Bits	Description
15:00	**Temperature Sensor Temperature (TST):** Contains the current temperature in degrees Kelvin reported by the temperature sensor. The physical point in the NVM subsystem whose temperature is reported by the temperature sensor and the temperature accuracy is implementation specific. An implementation that does not implement the temperature sensor reports a temperature of zero degrees Kelvin. The temperature reported by a temperature sensor may be used to trigger an asynchronous event

5.6 SET Features Identifiers

In addition to the SMART attribute structure, features pertaining to the operation and health of the D5-P4420 can be reported to the host on request through the Get Features command. Get Features support is per device and not available for a per namespace basis. The user can change settings using SET Features on the following items as defined in NVMe 1.2.1 specification.

Please refer to NVMe 1.2.1 specification for details of the following SET features:

- Arbitration (Feature Identifier 01h)
- Power Management (Feature Identifier 02h)
- Temperature Threshold (Feature Identifier 04h)
- Error Recovery (Feature Identifier 05h)
- Volatile Write Cache (Feature Identifier 06h)
- Number of Queues (Feature Identifier 07h)
- Interrupt Coalescing (Feature Identifier 08h)
- Interrupt Vector Configuration (Feature Identifier 09h)
- Write Atomicity (Feature Identifier 0Ah)
- Asynchronous Event Configuration (Feature Identifier 0Bh)

D5-P4420 also supports the following Vendor Unique Opcodes:

- Get Intel Log (Opcode D2h)

5.7 Vendor Unique Opcodes

5.7.1 Get Intel Log (Opcode D2h)

This vendor-unique Opcode helps read NLOGs (internal debug logs) from the drive. Get Log Page support is per device and not available for a per namespace basis.

Table 19: Get Intel Log - PRP Entry 1

Bit	Description
63:00	PRP Entry 1 (PRP1): This field contains the first PRP entry, specifying the start of the data buffer. If no data structure is used as part of the specified feature, then this field is ignored.

Table 20: Get Intel Log - PRP Entry 2

Bit	Description
63:00	PRP Entry 2 (PRP2): This field contains the second PRP entry. If the data transfer is satisfied with PRP Entry 1, then this field is reserved. If the data transfer may be satisfied with two PRP entries total, then this entry specifies the location where data should be transferred to. If the data transfer requires more than two PRP entries, then this field includes a pointer to a PRP List

Table 21: Get Intel Log - Command DWORD 10

Bit	Description
21:00	Num DWORDS. This indicates the number of Dwords to transfer for the segment of the Log being read

Table 22: Get Intel Log - Command DWORD 12

Bit	Description
63:00	Reserved
02:00	Select Log: Write Usage: 000 = Nlog, 001 = Event Log, 010 = Assert Dump, 011 = Reserved; 100 = Bridge Chip Nlog

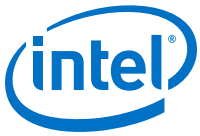
Table 23: Get Intel Log - Command DWORD 13

Bit	Description
31:00	Starting Offset: This field indicates the offset to start of the log

5.8 Vendor Unique FIDs

Intel® SSD D5-P4420 also supports the following Vendor Unique FIDs (VU) SET Features.

- Set/Get Max LBA (Feature Identifier C1h)
- Set/Get Native Max LBA (Feature Identifier C2h)
- Power Governor Setting (Feature Identifier C6h)
- Get/Set SMB ASIC Address (Feature Identifier C8h)
- Set/Get Blink activity for LED (Feature Identifier C9h)



- Reset Timed Workload Counters (Feature Identifier D5h)
- Set/Get Enable Latency Tracking (Feature Identifier E2h)

Table 24: Set Max LBA Setting - Command Dword 11 and Command Dword 12

Bit	Description
63:00	Maximum User LBA: Write Usage: This field sets the 64-bit maximum LBA addressable by the Drive. Read Usage: This field contains the 64-bit maximum LBA addressable by the Drive. Command Dword 11 contains bits 31:00; Command Dword 12 contains bits 63: 32.

Table 25: Status Code - Set Max LBA Command Specific Status Values

Value	Description
C0h	Requested MAX LBA exceeds Available capacity
C1h	Requested MAX LBA smaller than minimum allowable limit.
C2h	Requested MAX LBA is smaller than allocated Namespace requirements

Table 26: C6h - Set/Get Power (Typical) Governor Setting - Command Dword 11

Bit	Description
31:08	Reserved
07:00	Power Governor Setting: 00h = 25W, 01h = 20W, 02h = 15W

Table 27: Status Codes - Power Governor Setting Command Specific Status Value

Value	Description
C0h	Invalid Setting

Table 28: D5h - Reset Timed Workload Counters - Command Dword 11

Bit	Description
31:01	Reserved
00	Timed Workload Reset Settings: Write Usage: 00 = NOP, 1 = Reset E2, E3,E4 counters; Read Usage: Not Supported

Note: Get Features will not work for "Reset Timed Workload Counters," and status code is same as in the table above.

Table 29: E2h - Set/Get Enable Latency Tracking

Bit	Description
31:01	Write Usage: 00h = Disable Latency Tracking (Default) 01h = Enable Latency Tracking

Table 30: C8h - Get/Set SMB ASIC Address

Bit	Description
0	Reserved
09:01	SMB Controller Address
31:10	Reserved

Table 31: C9h - Set/Get Blink activity for LED

Feature Option	Feature Value Range	Description	Default
0	0-1	LED state while host is inactive. 0=off, 1=on	1 (on)
1	0-1	LED duration increment size. 0=50ms, 1=25ms	0 (50ms)
2	0-15	Off duration during IO activity in 25ms/ 50ms increments. 0=solid on (*)	0 (on)
3	0-15	On duration during IO activity in 25ms/ 50ms increments. 0=match off duration	0
4	0-15	Off duration during format activity in 25/50ms increments. 0=solid on (**)	5 (250ms)
5	0-15	On duration during format activity in 25/50ms increments. 0=match off duration	5 (250ms)

Notes:

- C9- Set Features Command Dword 11 will be divided into following sections
  - Bits[31:24] - Feature options and Bits[23:0] - Feature Value
  - Feature options and value ranges are defined above
- C9- Get Features Command Dword 11 will be divided into following sections
  - Bits[31:24] - Feature options and Bits[23:0] - Reserved. Must be 0.
  - Current value for the requested LED feature option will be returned in DW0[31:0]

§

6 NVMe Driver Support

The following table describes the NVMe Driver Support for D5-P4420. The support includes releasing and validating NVMe drivers for certain operating systems and validating functionality for open source drive, inbox, or native drivers for select operating systems.

Table 32: D5-P4420 NVMe Driver Support

Support Level	Operating System Description
Intel Provided[1]	Windows* Server 2016, Windows* Server 2012 R2, Win PE 10
In-box Driver or external package[2]	RHEL 7.4, 7.3, 7.2 CentOS 7.5, 7.3

Notes:

1. With Intel provided driver, full product specification is guaranteed, booting is only supported for 64bit OS.
2. With open source non-Intel driver, compatibility and functionality is validated.

§

7 Other Compliance and Certifications

Table 33: Other Compliance and Certifications

PCIe	Indicates compliance with PCI-SIG Organization testing requirements.
NVMe	Indicates compliance with UNH-IOL testing for NVMe compliance

§

Appendices

A Performance and Endurance Metrics

Table 34: User Addressable Sectors (LBA)

Capacity	Unformatted Capacity[1] (Total User Addressable Sectors in LBA Mode)
7.68TB [2]	15,002,931,888

Notes:

1. The total usable capacity of the SSD may be less than the total physical capacity because a small portion of the capacity is used for NAND media management and maintenance. IDEMA or JEDEC standard is used.
2. 1TB = 10^{12} bytes; 1 sector = 512 bytes LBA count shown represents total user storage capacity and will remain the same throughout the life of the drive.

Table 35: Random Read/Write Performance (IOPS)[1]

Intel® SSD D5-P4420	Queue Depth	Workers	Pre-conditioned with random write pattern	Pre-conditioned with sequential write pattern
Random 4KB[2] Write	64	4	46,000	46,000
Random 8KB[2] Write	64	4	23,000	23,000
Random 4KB Read	64	4	427,000	427,000
Random 8KB Read	64	4	240,000[3]	302,000[4]
Random 4KB 70/30 Read/Write	64	4	108,000	108,000
Random 8KB 70/30 Read/Write	64	4	55,000	55,000

Notes:

1. Performance measured using FIO3.5, CentOS 7.5.1804, Kernel: 4.4.74. Measurements are performed on a full Logical Block Address (LBA) span of the drive. Power mode set at PM0.
2. 4KB = 4,096 bytes; 8KB = 8,192 bytes
3. The drive is pre-conditioned with random write data pattern before measuring the random read.
4. The drive is pre-conditioned with sequential write data pattern before measuring the random read. The performance of large random read such as 8KB and16KB is improved by 20%. Intel recommends the workload of sequential write and random read to achieve the best performance

Table 36: Read/Write Consistency (%)[1]

Intel® SSD D5-P4420	Queue Depth	Workers	Consistency
Random 4KB Read [2]	32	4	90
Random 8KB Read	32	4	90
Random 4KB Write	32	4	70
Random 8KB Write	32	4	70
Sequential 128KB Write	128	1	75
Sequential 128KB Write	128	1	75

Notes:

1. Performance measured using FIO3.5, CentOS 7.5.1804, Kernel 4.14.74. Measurements are performed on a full Logical Block Address (LBA) span of the drive. Power mode set at PM0.
2. 4KB = 4,096 bytes; 8KB = 8,192 bytes; 128KB = 131,072 bytes.

Table 37: Sequential Read and Write Bandwidth (MB/s)[1]

Intel® SSD D5-P4420	
Sequential Read	**Sequential Write**
3,300	1,000

Note:

1. Performance measured using FIO3.5, CentOS 7.5.1804, Kernel 4.14.74 with 128KB (131,072 bytes) of transfer size and Queue Depth 32 and 4 workers (QD126). Measurements are performed on a full Logical Block Address (LBA) span of the drive. Power mode set at PM0

Table 38: Latency[1]

Intel® SSD D5-P4420		
Read[2]	**Write[3]**	**Power On to Controller Ready[4]**
131µs	25µs	<20s

Notes:

1. Performance measured using FIO3.5, CentOS 7.5.1804, Kernel 4.14.74.
2. Read latency is measured using 4 KB (4,096 bytes) transfer size with Queue Depth equal to 1 on the random workload.
3. Write latency is measured using 4 KB (4,096 bytes) transfer size with Queue Depth equal to 1 on the random workload.
4. Power On to Controller Ready (TTR) signifies when drive can begin receiving PCIe commands from host based on a single #PERESET. For power on from unsafe shutdown, power on to controller ready can take up to 60 seconds.

Table 39: Quality of Service (QoS)[1] (µs)

Intel® SSD D5-P4420	
4K Random Read	**8K Random Read**
Quality of Service [2,3] (99%)	
230	318
Quality of Service [2,3] (99.99%)	
825	1060

Notes:

1. Performance measured using FIO3.5, CentOS 7.5.1804, Kernel 4.14.74. Power mode set at PM0.
2. Measured using random workload with Queue Depth equal to 1 as the time taken for 99 (or 99.99) percentile of commands to finish the round-trip from host to drive and back to the host.
3. Random read result is based on the drive pre-conditioned with sequential write pattern

Table 40: Endurance (DWPD)

Intel® SSD D5-P4420			
JESD219A		**64K Sequential Write**	
3 Year Warranty	**5 Year Warranty**	**3 Year Warranty**	**5 Year Warranty**
0.66	0.4	2.92	1.76

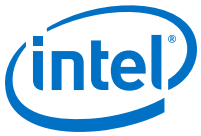


Table 41: Endurance (PBW)[1]

Intel® SSD D5-P4420	
JESD219A	**64K Sequential Write**
5 .6	24.6

Note: 1. 1PB = 10^{15} bytes

§

B Power Metrics

Table 42: Power Consumption (W)

Specification	Intel SSD D5-P4220
Average Active Write Power[1]	15
Average Active Read Power[2]	10
Max Burst Power[3]	19
Idle	~5

Notes:

1. The workload equates QD256/128KB Sequential Writes. Average power is measured over a 100ms sample period
2. The workload equates QD256/128KB Sequential Read. Average power is measured over a 100ms sample period
3. The workload equates QD256/128KB Sequential Writes. Max burst power is measured over a 500us sample period.

§

C IDENTIFY Data Structure

Table 43: Identify Controller

Bytes	O/M	Title	Description	Expected Value
Controller Capabilities and Features				
1:00	M	PCI Vendor ID (VID)	Contains the company vendor identifier that is assigned by the PCI SIG. This is the same value as reported in the ID register in section 2.1.1.	0x8086
3:02	M	PCI Subsystem Vendor ID (SSVID)	Contains the company vendor identifier that is assigned by the PCI SIG for the subsystem.	0x8086
23:04	M	Serial Number (SN)	Contains the serial number for the NVM subsystem that is assigned by the vendor as an ASCII string.	<variable>
63:24	M	Model Number (MN)	Contains the model number for the NVM subsystem that is assigned by the vendor as an ASCII string.	<variable>
71:64	M	Firmware Revision (FR)	Contains the currently active firmware revision for the NVM subsystem. This is the same revision information that may be retrieved with the Get Log Page command.	<variable>
72	M	Recommended Arbitration Burst (RAB)	This is the recommended Arbitration Burst size. The value is in commands and is reported as a power of two (2^n). This is the same units as the Arbitration Burst size.	0x00
75:73	M	IEEE OUI Identifier (IEEE)	Contains the Organization Unique Identifier (OUI) for the controller vendor. The OUI shall be a valid IEEE/RAC assigned identifier that may be registered at http://standards.ieee.org/develop/regauth/oui/public.html.	0x5CD2E4
76	O	Controller Multi-Path I/O and Namespace Sharing Capabilities (CMIC)	This field specifies multi-path I/O and namespace sharing capabilities of the controller and NVM subsystem.	0x00
			Bits 7:3 are reserved	
			Bit 2: If set to '1' then the controller is associated with an SR-IOV Virtual Function. If cleared to '0' then the controller is associated with a PCI Function.	0
			Bit 1: If set to '1' then the NVM subsystem may contain two or more controllers. If cleared to '0' then the NVM subsystem contains only a single controller.	0
			Bit 0: If set to '1' then the NVM subsystem may contain two or more physical PCI Express ports. If cleared to '0' then the NVM subsystem contains only a single PCI Express port.	0
77	M	Maximum Data Transfer Size (MDTS)	This field indicates the maximum data transfer size between the host and the controller. The host should not issue a command that exceeds this transfer size. If a command is processed that exceeds the transfer size, then the command is aborted with a status of Invalid Field in Command. The value is in units of the minimum memory page size (CAP.MPSMIN) and is reported as a power of two (2^*n*). A value of 0h indicates no restrictions on transfer size. The restriction includes metadata if it is interleaved with the logical block data.	0x05 (128KB)
79:78	M	Controller ID (CNTLID)	Contains the NVM subsystem unique controller identifier associated with the controller.	0x00
83:80	M	Version (VER)	This field contains the value reported in the Version register defined in section 3.1.2. Implementations compliant to revision 1.2 or later of this specification shall report a non-zero value in this field.	0x00 01 02 00
87:84	M	RTD3 Resume Latency (RTD3R)	This field indicates the typical latency in microseconds resuming from Runtime D3 (RTD3). A value of 0h indicates RTD3 Resume Latency is not reported.	0x01312D00

continued...

Bytes	O/M	Title	Description	Expected Value
91:88	M	RTD3 Entry Latency (RTD3E)	This field indicates the typical latency in microseconds to enter Runtime D3 (RTD3). A value of 0h indicates RTD3 Entry Latency is not reported.	0x00989680
95:92	M	Optional Asynchronous Events Supported (OAES)	This field indicates the optional asynchronous events supported by the controller. A controller shall not send optional asynchronous events before they are enabled by host software.	0x0
			Bits 31:10 are reserved.	0
			Bit 9 is set to '1' if the controller supports sending Firmware Activation Notices. If cleared to '0' then the controller does not support the Firmware Activation Notices event.	0
			Bit 8 is set to '1' if the controller supports sending the Namespace Attribute Changed Notices. If cleared to '0' then the controller does not support the Namespace Attribute Notices event.	0
			Bits 7:0 are reserved.	0
239:96			Reserved	
255:240			Reserved	
Admin Command Set Attributes				
257:256	M	Optional Admin Command Support (OACS):	This field indicates the optional Admin commands supported by the controller.	0x6
			Bits 15:4 are reserved.	
			Bit 3 if set to '1' then the controller supports the Namespace Management and Namespace Attachment commands. If cleared to '0' then the controller does not support the Namespace Management and Namespace Attachment commands.	0
			Bit 2 if set to '1' then the controller supports the Firmware Activate and Firmware Download commands. If cleared to '0' then the controller does not support the Firmware Activate and Firmware Download commands.	1
			Bit 1 if set to '1' then the controller supports the Format NVM command. If cleared to '0' then the controller does not support the Format NVM command.	1
			Bit 0 if set to '1' then the controller supports the Security Send and Security Receive commands. If cleared to '0' then the controller does not support the Security Send and Security Receive commands.	0
258	M	Abort Command Limit (ACL)	This field is used to convey the maximum number of concurrently outstanding Abort commands supported by the controller. This is a 0's based value. It is recommended that implementations support a minimum of four Abort commands outstanding simultaneously.	0x3 (4)
259	M	Asynchronous Event Request Limit (AERL)	This field is used to convey the maximum number of concurrently outstanding Asynchronous Event Request commands supported by the controller. This is a 0's based value. It is recommended that implementations support a minimum of four Asynchronous Event Request Limit commands outstanding simultaneously.	0x03 (4)
260	M	Firmware Updates (FRMW)	This field indicates capabilities regarding firmware updates.	0x02
			Bits 7:5 are reserved.	
			Bit 4 if set to '1' indicates that the controller supports firmware activation without a reset. If cleared to '0' then the controller requires a reset for firmware to be activated.	0

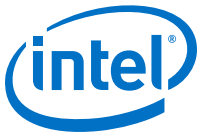
continued...

Bytes	O/M	Title	Description	Expected Value
			Bits 3:1 indicate the number of firmware slots that the device supports. This field shall specify a value between one and seven, indicating that at least one firmware slot is supported and up to seven maximum. This corresponds to firmware slots 1 through 7.	1
			Bit 0 if set to '1' indicates that the first firmware slot (slot 1) is read only. If cleared to '0' then the first firmware slot (slot 1) is read/write. Implementations may choose to have a baseline read only firmware image.	0
261	M	Log Page Attributes (LPA)	This field indicates optional attributes for log pages that are accessed via the Get Log Page command.	0x02
			Bits 7:2 are reserved.	
			Bit 1 if set to '1' then the controller supports the Command Effects log page. Bit 1 if cleared to '0' then the controller does not support the Command Effects log page.	1
			Bit 0 if set to '1' then the controller supports the SMART / Health information log page on a per namespace basis. If cleared to '0' then the controller does not support the SMART / Health information log page on a per namespace basis; the log page returned is global for all namespaces.	0
262	M	Error Log Page Entries (ELPE)	This field indicates the maximum number of Error Information log entries that are stored by the controller. This field is a 0's based value.	0x3F (64)
263	M	Number of Power States Support (NPSS)	This field indicates the number of NVM Express power states supported by the controller. This is a 0's based value.	0x00
			Power states are numbered sequentially starting at power state 0. A controller shall support at least one power state (i.e., power state 0) and may support up to 31 additional power states (i.e., up to 32 total).	
264	M	Admin Vendor Specific Command Configuration (AVSCC)	This field indicates the configuration settings for Admin Vendor Specific command handling.	0x00
			Bits 7:1 are reserved.	
			Bit 0 if set to '1' indicates that all Admin Vendor Specific Commands use the format defined in Figure 8 of NVMe Spec 1.2. If cleared to '0' indicates that the format of all Admin Vendor Specific Commands are vendor specific.	
265	O	Autonomous Power State Transition Attributes (APSTA)	This field indicates the attributes of the autonomous power state transition feature. Bits 7:1 are reserved. Bit 0 if set to '1' then the controller supports autonomous power state transitions. If cleared to '0' then the controller does not support autonomous power state transitions.	0x00
267:266	M	Warning Composite Temperature Threshold (WCTEMP)	This field indicates the minimum Composite Temperature field value (reported in the SMART / Health Information log that indicates an overheating condition during which controller operation continues. Immediate remediation is recommended (e.g., additional cooling or workload reduction). The platform should strive to maintain a composite temperature below this value. A value of 0h in this field indicates that no warning temperature threshold value is reported by the controller. Implementations compliant to revision 1.2 or later of this specification shall report a non-zero value in this field. It is recommended that implementations report a value of 0157h in this field.	0x0157 (343 Kelvin)

continued...

Bytes	O/M	Title	Description	Expected Value
269:268	M	Critical Composite Temperature Threshold (CCTEMP)	This field indicates the minimum Composite Temperature field value (reported in the SMART / Health Information log) that indicates a critical overheating condition (e.g., may prevent continued normal operation, possibility of data loss, automatic device shutdown, extreme performance throttling, or permanent damage). A value of 0h in this field indicates that no critical temperature threshold value is reported by the controller. Implementations compliant to revision 1.2 or later of this specification shall report a non-zero value in this field.	0x161 (353 Kelvin)
271:270	O	Maximum Time for Firmware Activation (MTFA)	Indicates the maximum time the controller temporarily stops processing commands to activate the firmware image. This field shall be valid if the controller supports firmware activation without a reset. This field is specified in 100 millisecond units. A value of 0h indicates that the maximum time is undefined.	0x0000
275:272	O	Host Memory Buffer Preferred Size (HMPRE)	This field indicates the preferred size that the host is requested to allocate for the Host Memory Buffer feature in 4KB units. This value shall be larger than or equal to the Host Memory Buffer Minimum Size. If this field is non-zero, then the Host Memory Buffer feature is supported. If this field is cleared to 0h, then the Host Memory Buffer feature is not supported.	0x00
279:276	O	Host Memory Buffer Minimum Size (HMMIN)	This field indicates the minimum size that the host is requested to allocate for the Host Memory Buffer feature in 4KB units. If this field is cleared to 0h, then the host is requested to allocate any amount of host memory possible up to the HMPRE value.	0x00
295:280	O	Total NVM Capacity (TNVMCAP)	This field indicates the total NVM capacity in the NVM subsystem. The value is in bytes. This field shall be supported if Namespace Management and Namespace Attachment commands are supported.	N/A
311:296	O	Unallocated NVM Capacity (UNVMCAP)	This field indicates the unallocated NVM capacity in the NVM subsystem. The value is in bytes. This field shall be supported if Namespace Management and Namespace Attachment commands are supported.	N/A
315:312	O	Replay Protected Memory Block Support (RPMBS)	This field indicates if the controller supports one or more Replay Protected Memory Blocks (RPMBs) and the capabilities.	0x00
511:316			Reserved	
NVM Command Set Attributes				
512	M	Submission Queue Entry Size (SQES)	This field defines the required and maximum Submission Queue entry size when using the NVM Command Set.	0x66
			Bits 7:4 define the maximum Submission Queue entry size when using the NVM Command Set. This value is larger than or equal to the required SQ entry size. The value is in bytes and is reported as a power of two (2^*n*). The recommended value is 6, corresponding to a standard NVM Command Set SQ entry size of 64 bytes. Controllers that implement proprietary extensions may support a larger value.	
			Bits 3:0 define the required Submission Queue Entry size when using the NVM Command Set. This is the minimum entry size that may be used. The value is in bytes and is reported as a power of two (2^*n*). The required value shall be 6, corresponding to 64.	
513	M	Completion Queue Entry Size (CQES)	This field defines the required and maximum Completion Queue entry size when using the NVM Command Set.	0x44
			Bits 7:4 define the maximum Completion Queue entry size when using the NVM Command Set. This value is larger than or equal to the required CQ entry size. The value is in bytes and is reported as a power of two (2^*n*). The recommended value is 4, corresponding to a standard NVM Command Set CQ entry size of 16 bytes. Controllers that implement proprietary extensions may support a larger value.	

continued...



Bytes	O/M	Title	Description	Expected Value
			Bits 3:0 define the required Completion Queue entry size when using the NVM Command Set. This is the minimum entry size that may be used. The value is in bytes and is reported as a power of two (2^*n*). The required value shall be 4, corresponding to 16.	
515:514			Reserved	
519:516	M	Number of Namespaces (NN)	This field defines the number of valid namespaces present for the controller. Namespaces shall be allocated in order (starting with 1) and packed sequentially.	0x01
521:520	M	Optional NVM Command Support (ONCS)	This field indicates the optional NVM commands supported by the controller.	0x06
			Bits 15:6 are reserved.	
			Bit 5 if set to '1' then the controller supports reservations. If cleared to '0' then the controller does not support reservations. If the controller supports reservations, then it shall support the following commands associated with reservations: Reservation Report, Reservation Register, Reservation Acquire, and Reservation Release.	0
			Bit 4 if set to '1' then the controller supports the Save field in the Set Features command and the Select field in the Get Features command. If cleared to '0' then the controller does not support the Save field in the Set Features command and the Select field in the Get Features command.	0
			Bit 3 if set to '1' then the controller supports the Write Zeroes command. If cleared to '0' then the controller does not support the Write Zeroes command.	0
			Bit 2 if set to '1' then the controller supports the Dataset Management command. If cleared to '0' then the controller does not support the Dataset Management command.	1
			Bit 1 if set to '1' then the controller supports the Write Uncorrectable command. If cleared to '0' then the controller does not support the Write Uncorrectable command.	1
			Bit 0 if set to '1' then the controller supports the Compare command. If cleared to '0' then the controller does not support the Compare command.	0
523:522	M	Fused Operation Support (FUSES)	This field indicates the fused operations that the controller supports.	0x00
			Bits 15:1 are reserved.	
			Bit 0 if set to '1' then the controller supports the Compare and Write fused operation. If cleared to '0' then the controller does not support the Compare and Write fused operation. Compare shall be the first command in the sequence.	
524	M	Format NVM Attributes (FNA)	This field indicates attributes for the Format NVM command.	0x04
			Bits 7:3 are reserved.	
			Bit 2 indicates whether cryptographic erase is supported as part of the secure erase functionality. If set to '1', then cryptographic erase is supported. If cleared to '0', then cryptographic erase is not supported.	1
			Bit 1 indicates whether secure erase functionality applies to all namespaces or is specific to a particular namespace. If set to'1', then a secure erase of a particular namespace as part of a format results in a secure erase of all namespaces. If cleared to '0', then a secure erase as part of a format is performed on a per namespace basis.	N/A

continued...

Bytes	O/M	Title	Description	Expected Value
			Bit 0 indicates whether the format operation applies to all namespaces or is specific to a particular namespace. If set to '1', then all namespaces shall be configured with the same attributes and a format of any namespace results in a format of all namespaces. If cleared to '0', then the controller supports format on a per namespace basis.	N/A
525	M	Volatile Write Cache (VWC)	This field indicates attributes related to the presence of a volatile write cache in the implementation.	0x00
			Bits 7:1 are reserved.	
			Bit 0 if set to '1' indicates that a volatile write cache is present. If cleared to '0', a volatile write cache is not present. If a volatile write cache is present, then the host may issue Flush commands and control whether it is enabled with Set Features specifying the Volatile Write Cache feature identifier. If a volatile write cache is not present, the host shall not issue Flush commands nor Set Features or Get Features with the Volatile Write Cache identifier.	
527:526	M	Atomic Write Unit Normal (AWUN)	This field indicates the atomic write size for the controller during normal operation. This field is specified in logical blocks and is a 0's based value. If a write is issued of this size or less, the host is guaranteed that the write is atomic to the NVM with respect to other read or write operations. A value of FFFFh indicates all commands are atomic as this is the largest command size. It is recommended that implementations support a minimum of 128KB (appropriately scaled based on LBA size).	0x00
529:528	M	Atomic Write Unit Power Fail (AWUPF)	This field indicates the atomic write size for the controller during a power fail condition. This field is specified in logical blocks and is a 0's based value. If a write is issued of this size or less, the host is guaranteed that the write is atomic to the NVM with respect to other read or write operations.	0x00
530	M	NVM Vendor Specific Command Configuration (NVSCC)	This field indicates the configuration settings for NVM Vendor Specific command handling.	0x00
			Bits 7:1 are reserved.	
			Bit 0 if set to '1' indicates that all NVM Vendor Specific Commands use the format defined in Figure 8. If cleared to '0' indicates that the format of all NVM Vendor Specific Commands are vendor specific.	0
531	M		Reserved	
533:532	O	Atomic Compare & Write Unit (ACWU)	This field indicates the size of the write operation guaranteed to be written atomically to the NVM across all namespaces with any supported namespace format for a Compare and Write fused operation. If a specific namespace guarantees a larger size than is reported in this field, then this namespace specific size is reported in the NACWU field in the Identify Namespace data structure. This field shall be supported if the Compare and Write fused command is supported. This field is specified in logical blocks and is a 0's based value. If a Compare and Write is submitted that requests a transfer size larger than this value, then the controller may fail the command with a status code of Invalid Field in Command. If Compare and Write is not a supported fused command, then this field shall be 0h.	0x00
535:534	M		Reserved	
539:536	O	SGL Support (SGLS)	This field indicates if SGLs are supported for the NVM Command Set and the particular SGL types supported.	0x00
			Bits[31:20] Reserved	

continued...

Bytes	O/M	Title	Description	Expected Value
			Bit 19: If set to '1', then use of a Metadata Pointer (MPTR) that contains an address of an SGL segment containing exactly one SGL Descriptor that is Qword aligned is supported. If cleared to '0', then use of a MPTR containing an SGL Descriptor is not supported.	0
			Bit 18 If set to '1', then the controller supports commands that contain a data or metadata SGL of a length larger than the amount of data to be transferred. If cleared to '0', then the SGL length shall be equal to the amount of data to be transferred.	0
			Bit 17 If set to '1', then use of a byte aligned contiguous physical buffer of metadata (the Metadata Pointer field in Figure 12) is supported. If cleared to '0', then use of a byte aligned contiguous physical buffer of metadata is not supported.	0
			Bit 16 If set to '1', then the SGL Bit Bucket descriptor is supported. If cleared to '0', then the SGL Bit Bucket descriptor is not supported.	0
			Bit[15:1] Reserved	
			Bit 0 If set to '1', then the controller supports SGLs for the NVM Command Set including the SGL Data Block, SGL Segment, and SGL Last Segment descriptor types. If cleared to '0', then the controller does not support SGLs for the NVM Command Set and all other bits in this field shall be cleared to '0'.	0
703:540	M		Reserved	
I/O Command Set Attributes				
2047:704			Reserved	
Power State Descriptors				
2079:2048	M	Power State 0 Descriptor (PSD0)	This field indicates the characteristics of power state 0. Table 44 on page 49	Data Structure in Table 44 on page 49
2111:2080	O	Power State 1 Descriptor (PSD1)	This field indicates the characteristics of power state 1. Table 44 on page 49	
2143:2112	O	Power State 2 Descriptor (PSD2)	This field indicates the characteristics of power state 2. Table 44 on page 49	
2175:2144	O	Power State 3 Descriptor (PSD3)	This field indicates the characteristics of power state 3. Table 44 on page 49	
2207:2176	O	Power State 4 Descriptor (PSD4)	This field indicates the characteristics of power state 4. Table 44 on page 49	
2239:2208	O	Power State 5 Descriptor (PSD5)	This field indicates the characteristics of power state 5. Table 44 on page 49	
2271:2240	O	Power State 6 Descriptor (PSD6)	This field indicates the characteristics of power state 6. Table 44 on page 49	
2303:2272	O	Power State 7 Descriptor (PSD7)	This field indicates the characteristics of power state 7. Table 44 on page 49	
2335:2304	O	Power State 8 Descriptor (PSD8)	This field indicates the characteristics of power state 8. Table 44 on page 49	
2367:2336	O	Power State 9 Descriptor (PSD9)	This field indicates the characteristics of power state 9. Table 44 on page 49	
2399:2368	O	Power State 10 Descriptor (PSD10)	This field indicates the characteristics of power state 10. Table 44 on page 49	

continued...

Bytes	O/M	Title	Description	Expected Value
2431:2400	O	Power State 11 Descriptor (PSD11)	This field indicates the characteristics of power state 11. Table 44 on page 49	
2463:2432	O	Power State 12 Descriptor (PSD12)	This field indicates the characteristics of power state 12. Table 44 on page 49	
2495:2464	O	Power State 13 Descriptor (PSD13)	This field indicates the characteristics of power state 13. Table 44 on page 49	
2527:2496	O	Power State 14 Descriptor (PSD14)	This field indicates the characteristics of power state 14. Table 44 on page 49	
2559:2528	O	Power State 15 Descriptor (PSD15)	This field indicates the characteristics of power state 15. Table 44 on page 49	
2591:2560	O	Power State 16 Descriptor (PSD16)	This field indicates the characteristics of power state 16. Table 44 on page 49	
2623:2592	O	Power State 17 Descriptor (PSD17)	This field indicates the characteristics of power state 17. Table 44 on page 49	
2655:2624	O	Power State 18 Descriptor (PSD18)	This field indicates the characteristics of power state 18. Table 44 on page 49	
2687:2656	O	Power State 19 Descriptor (PSD19)	This field indicates the characteristics of power state 19. Table 44 on page 49	
2719:2688	O	Power State 20 Descriptor (PSD20)	This field indicates the characteristics of power state 20. Table 44 on page 49	
2751:2720	O	Power State 21 Descriptor (PSD21):	This field indicates the characteristics of power state 21. Table 44 on page 49	
2783:2752	O	Power State 22 Descriptor (PSD22)	This field indicates the characteristics of power state 22. Table 44 on page 49	
2815:2784	O	Power State 23 Descriptor (PSD23)	This field indicates the characteristics of power state 23. Table 44 on page 49	
2847:2816	O	Power State 24 Descriptor (PSD24)	This field indicates the characteristics of power state 24. Table 44 on page 49	
2879:2848	O	Power State 25 Descriptor (PSD25)	This field indicates the characteristics of power state 25. Table 44 on page 49	
2911:2880	O	Power State 26 Descriptor (PSD26)	This field indicates the characteristics of power state 26. Table 44 on page 49	
2943:2912	O	Power State 27 Descriptor (PSD27)	This field indicates the characteristics of power state 27. Table 44 on page 49	
2975:2944	O	Power State 28 Descriptor (PSD28)	This field indicates the characteristics of power state 28. Table 44 on page 49	
3007:2976	O	Power State 29 Descriptor (PSD29)	This field indicates the characteristics of power state 29. Table 44 on page 49	
3039:3008	O	Power State 30 Descriptor (PSD30)	This field indicates the characteristics of power state 30. Table 44 on page 49	
3071:3040	O	Power State 31 Descriptor (PSD31)	This field indicates the characteristics of power state 31. Table 44 on page 49	

Table 44: Power State Descriptors Data Structure

Bytes	Description	Expected Value
255:184	Reserved	
183:182	Active Power Scale (APS): This field indicates the scale for the Active Power field. If an Active Power Workload is reported for a power state, then the Active Power Scale shall also be reported for that power state.	0x00
181:179	Reserved	
178:176	Active Power Workload (APW): This field indicates the workload used to calculate maximum power for this power state. This field shall not be “No Workload” unless ACTP is 0000h.	0x00
175:160	Active Power (ACTP): This field indicates the largest average power consumed by the NVM subsystem over a 10 second period in this power state with the workload indicated in the Active Power Workload field. The power in Watts is equal to the value in this field multiplied by the scale indicated in the Active Power Scale field. A value of 0000h indicates Active Power is not reported.	0x00
159:152	Reserved	
151:150	Idle Power Scale (IPS): This field indicates the scale for the Idle Power field.	0x00
149:144	Reserved	
143:128	Idle Power (IDLP): This field indicates the typical power consumed by the NVM subsystem over 30 seconds in this power state when idle (i.e., there are no pending commands, register accesses, nor background processes). The measurement starts after the NVM subsystem has been idle for 10 seconds. The power in Watts is equal to the value in this field multiplied by the scale indicated in the Idle Power Scale field. A value of 0000h indicates Idle Power is not reported.	0x00
127:125	Reserved	
124:120	Relative Write Latency (RWL): This field indicates the relative write latency associated with this power state. The value in this field shall be less than the number of supported power states (e.g., if the controller supports 16 power states, then valid values are 0 through 15). A lower value means lower write latency.	0x00
119:117	Reserved	
116:112	Relative Write Throughput (RWT): This field indicates the relative write throughput associated with this power state. The value in this field shall be less than the number of supported power states (e.g., if the controller supports 16 power states, then valid values are 0 through 15). A lower value means higher write throughput.	0x00
111:109	Reserved	
108:104	Relative Read Latency (RRL): This field indicates the relative read latency associated with this power state. The value in this field shall be less than the number of supported power states (e.g., if the controller supports 16 power states, then valid values are 0 through 15). A lower value means lower read latency.	0x00
103:101	Reserved	
100:96	Relative Read Throughput (RRT): This field indicates the relative read throughput associated with this power state. The value in this field shall be less than the number of supported power states (e.g., if the controller supports 16 power states, then valid values are 0 through 15). A lower value means higher read throughput.	0x00
95:64	Exit Latency (EXLAT): This field indicates the maximum exit latency in microseconds associated with entering this power state.	0x00
63:32	Entry Latency (ENLAT): This field indicates the maximum entry latency in microseconds associated with entering this power state.	0x00
31:26	Reserved	

continued...

Bytes	Description	Expected Value
25	Non-Operational State (NOPS): This field indicates whether the controller processes I/O commands in this power state. If this field is cleared to '0', then the controller processes I/O commands in this power state. If this field is set to '1', then the controller does not process I/O commands in this power state.	0x00
24	Max Power Scale (MXPS): This field indicates the scale for the Maximum Power field. If this field is cleared to '0', then the scale of the Maximum Power field is in 0.01 Watts. If this field is set to '1', then the scale of the Maximum Power field is in 0.0001 Watts.	0x00
23:16	Reserved	
15:00	Maximum Power (MP): This field indicates the maximum power consumed by the NVM subsystem in this power state. The power in Watts is equal to the value in this field multiplied by the scale specified in the Max Power Scale field.	<variable>

Note: O = Optional. The content of the word is optional

M = Mandatory. The content of the word is mandatory

Table 45: Vendor Specific Data Structure

Vendor Specific			Expected Value
3072	Intel Vendor Specific	Vendor Specific (VS): This range of bytes is allocated for vendor specific usage.	
3073		Vendor Specific (VS): This range of bytes is allocated for vendor specific usage.	
3074		Vendor Specific (VS): This range of bytes is allocated for vendor specific usage.	
3075		Stripe size: Specifies the size of the stripe, value is read only and cannot be changed by SW. The value is in units of the minimum memory page size (CAP.MPSMIN) and is reported as a power of two (2^n) 0: Driver assisted striping not supported by FW 1: 8 KiB of user data (i.e. 16 512/520/528 byte or 2 4096 byte and extended sectors) 2 16 KiB of user data (i.e. 32 512/520/528 byte or 4 4096 byte and extended sectors) 3: 32 KiB of user data (i.e. 64 512/520/528 byte or 8 4096 byte and extended sectors) 4: 64 KiB of user data (i.e. 128 512/520/528 byte or 16 4096 byte and extended sectors) 5: 128 KiB of user data (i.e. 256 512/520/528 byte or 32 4096 byte and extended sectors)	0x05
3095:3076		Standardized Failure mode String	<variable>
3096		Current PCIe Link Speed field (CLS)	Healthy link would report Gen3 0x03
3097		Negotiated Link Width (NLW)	Healthy link would report 4 lanes 0x04
3098		Bit[31:0] Reserved	N/A
3099		Bits[31:0] Reserved	N/A
3107:3100		VS_BootloaderVersion	<variable per firmware version>
4095:3108	O	Vendor Specific (VS): This range of bytes is allocated for vendor specific usage	Reserved

Note: O = Optional. The content of the word is optional

M = Mandatory. The content of the word is mandatory

Table 46: Identify Namespace Data Structure

Identify Namespace Data Structure, NVM Command Set Specific			
Bytes	**O/M**	**Description**	**Expected Value**
7:00	M	Namespace Size (NSZE): This field indicates the total size of the namespace in logical blocks. A namespace of size *n* consists of LBA 0 through (*n* - 1). The number of logical blocks is based on the formatted LBA size. This field is undefined prior to the namespace being formatted.	Varies
15:08	M	Namespace Capacity (NCAP): This field indicates the maximum number of logical blocks that may be allocated in the namespace at any point in time. The number of logical blocks is based on the formatted LBA size. This field is undefined prior to the namespace being formatted. This field is used in the case of thin provisioning and reports a value that is smaller than or equal to the Namespace Size. Spare LBAs are not reported as part of this field.	Equal to NSZE if NS is formatted and attached to the controller; Else undefined.
		A logical block is allocated when it is written with a Write or Write Uncorrectable command. A logical block may be deallocated using the Dataset Management command.	
23:16	M	Namespace Utilization (NUSE): This field indicates the current number of logical blocks allocated in the namespace. This field is smaller than or equal to the Namespace Capacity. The number of logical blocks is based on the formatted LBA size.	Equal to NSZE
		When using the NVM command set: A logical block is allocated when it is written with a Write or Write Uncorrectable command. A logical block may be deallocated using the Dataset Management command.	
24	M	Namespace Features (NSFEAT): This field defines features of the namespace.	0x00
		Bits 7:3 are reserved.	
		Bit 2 if set to '1' indicates that the controller supports the Deallocated or Unwritten Logical Block error for this namespace. If cleared to '0', then the controller does not support the Deallocated or Unwritten Logical Block error for this namespace.	0
		Bit 1 if set to '1' indicates that the fields NAWUN, NAWUPF, and NACWU are defined for this namespace and should be used by the host for this namespace instead of the AWUN, AWUPF, and ACWU fields in the Identify Controller data structure. If cleared to '0', then the controller does not support the fields NAWUN, NAWUPF, and NACWU for this namespace. In this case, the host should use the AWUN, AWUPF, and ACWU fields defined in the Identify Controller data structure in Figure 90.	0
		Bit 0 if set to '1' indicates that the namespace supports thin provisioning. Specifically, the Namespace Capacity reported may be less than the Namespace Size. When this feature is supported and the Dataset Management command is supported then deallocating LBAs shall be reflected in the Namespace Utilization field. Bit 0 if cleared to '0' indicates that thin provisioning is not supported and the Namespace Size and Namespace Capacity fields report the same value.	0
25	M	Number of LBA Formats (NLBAF): This field defines the number of supported LBA size and metadata size combinations supported by the namespace. LBA formats shall be allocated in order (starting with 0) and packed sequentially. This is a 0's based value. The maximum number of LBA formats that may be indicated as supported is 16. The supported LBA formats are indicated in bytes 128 – 191 in this data structure.	0x1
		The metadata may be either transferred as part of the LBA (creating an extended LBA which is a larger LBA size that is exposed to the application) or it may be transferred as a separate contiguous buffer of data. The metadata shall not be split between the LBA and a separate metadata buffer.	
		It is recommended that software and controllers transition to an LBA size that is 4KB or larger for ECC efficiency at the controller. If providing metadata, it is recommended that at least 8 bytes are provided per logical block to enable use with end-to-end data protection.	
26	M	Formatted LBA Size (FLBAS): This field indicates the LBA size & metadata size combination that the namespace has been formatted with.	0x0 if 512B or 0x1 if 4096B
		Bits 7:5 are reserved.	0

continued...

Identify Namespace Data Structure, NVM Command Set Specific			
Bytes	**O/M**	**Description**	**Expected Value**
		Bit 4 if set to '1' indicates that the metadata is transferred at the end of the data LBA, creating an extended data LBA. Bit 4 if cleared to '0' indicates that all of the metadata for a command is transferred as a separate contiguous buffer of data.	0
		Bits 3:0 indicates one of the 16 supported combinations indicated in this data structure. This is a 0's based value.	0 or 1
27	M	Metadata Capabilities (MC): This field indicates the capabilities for metadata.	
		Bits 7:2 are reserved.	
		Bit 1 if set to '1' indicates the namespace supports the metadata being transferred as part of a separate buffer that is specified in the Metadata Pointer. Bit 1 if cleared to '0' indicates that the controller does not support the metadata being transferred as part of a separate buffer.	0
		Bit 0 if set to '1' indicates that the namespace supports the metadata being transferred as part of an extended data LBA. Specifically, the metadata is transferred as part of the data PRP Lists. Bit 0 if cleared to '0' indicates that the namespace does not support the metadata being transferred as part of an extended data LBA.	0
28	M	End-to-end Data Protection Capabilities (DPC): This field indicates the capabilities for the end-to-end data protection feature. Multiple bits may be set in this field.	0x00
		Bits 7:5 are reserved.	0
		Bit 4 if set to '1' indicates that the namespace supports protection information transferred as the last eight bytes of metadata. Bit 4 if cleared to '0' indicates that the namespace does not support protection information transferred as the last eight bytes of metadata.	0
		Bit 3 if set to '1' indicates that the namespace supports protection information transferred as the first eight bytes of metadata. Bit 3 if cleared to '0' indicates that the namespace does not support protection information transferred as the first eight bytes of metadata.	0
		Bit 2 if set to '1' indicates that the namespace supports Protection Information Type 3. Bit 2 if cleared to '0' indicates that the namespace does not support Protection Information Type 3.	0
		Bit 1 if set to '1' indicates that the namespace supports Protection Information Type 2. Bit 1 if cleared to '0' indicates that the namespace does not support Protection Information Type 2.	0
		Bit 0 if set to '1' indicates that the namespace supports Protection Information Type 1. Bit 0 if cleared to '0' indicates that the namespace does not support Protection Information Type 1.	0
29	M	End-to-end Data Protection Type Settings (DPS): This field indicates the Type settings for the end-to-end data protection feature.	0x00
		Bits 7:4 are reserved.	
		Bit 3 if set to '1' indicates that the protection information, if enabled, is transferred as the first eight bytes of metadata. Bit 3 if cleared to '0' indicates that the protection information, if enabled, is transferred as the last eight bytes of metadata.	0
		Bits 2:0 indicate whether Protection Information is enabled and the type of Protection Information enabled. The values for this field have the following meanings:	0
		Value	
		000b (Protection Information is not enabled)	
		001b (Protection Information is enabled, Type 1)	
		010b (Protection Information is enabled, Type 2)	

continued...

Identify Namespace Data Structure, NVM Command Set Specific			
Bytes	**O/M**	**Description**	**Expected Value**
		011b (Protection Information is enabled, Type 3)	
		100b – 111b (Reserved)	
30	O	Namespace Multi-path I/O and Namespace Sharing Capabilities (NMIC): This field specifies multi-path I/O and namespace sharing capabilities of the namespace. Bits 7:1 are reserved Bit 0: If set to '1' then the namespace may be accessible by two or more controllers in the NVM subsystem (i.e., may be a shared namespace). If cleared to '0' then the namespace is a private namespace and may only be accessed by the controller that returned this namespace data structure.	0x00
31	O	Reservation Capabilities (RESCAP): This field indicates the reservation capabilities of the namespace. A value of 00h in this field indicates that reservations are not supported by this namespace. Bit 7 is reserved Bit 6 if set to '1' indicates that the namespace supports the Exclusive Access – All Registrants reservation type. If this bit is cleared to '0', then the namespace does not support the Exclusive Access – All Registrants reservation type. Bit 5 if set to '1' indicates that the namespace supports the Write Exclusive – All Registrants reservation type. If this bit is cleared to '0', then the namespace does not support the Write Exclusive – All Registrants reservation type. Bit 4 if set to '1' indicates that the namespace supports the Exclusive Access – Registrants Only reservation type. If this bit is cleared to '0', then the namespace does not support the Exclusive Access – Registrants Only reservation type. Bit 3 if set to '1' indicates that the namespace supports the Write Exclusive – Registrants Only reservation type. If this bit is cleared to '0', then the namespace does not support the Write Exclusive – Registrants Only reservation type. Bit 2 if set to '1' indicates that the namespace supports the Exclusive Access reservation type. If this bit is cleared to '0', then the namespace does not support the Exclusive Access reservation type. Bit 1 if set to '1' indicates that the namespace supports the Write Exclusive reservation type. If this bit is cleared to '0', then the namespace does not support the Write Exclusive reservation type. Bit 0 if set to '1' indicates that the namespace supports the Persist Through Power Loss capability. If this bit is cleared to '0', then the namespace does not support the Persist Through Power Loss Capability.	0x00
32	O	Format Progress Indicator (FPI): If a format operation is in progress, this field indicates the percentage of the namespace that remains to be formatted. Bit 7 if set to '1' indicates that the namespace supports the Format Progress Indicator defined by bits 6:0 in this field. If this bit is cleared to '0', then the namespace does not support the Format Progress Indicator and bits 6:0 in this field shall be cleared to 0h. Bits 6:0 indicate the percentage of the namespace that remains to be formatted (e.g., a value of 25 indicates that 75% of the namespace has been formatted and 25% remains to be formatted). A value of 0 indicates that the namespace is formatted with the format specified by the FLBAS and DPS fields in this data structure.	0x00
33		Reserved	
35:34	O	Namespace Atomic Write Unit Normal (NAWUN): This field indicates the namespace specific size of the write operation guaranteed to be written atomically to the NVM during normal operation. A value of 0h indicates that the size for this namespace is the same size as that reported in the AWUN field of the Identify Controller data structure. All other values specify a size in terms of logical blocks using the same encoding as the AWUN field.	0x00
37:36	O	Namespace Atomic Write Unit Power Fail (NAWUPF): This field indicates the namespace specific size of the write operation guaranteed to be written atomically to the NVM during a power fail or error condition. A value of 0h indicates that the size for this namespace is the same size as that reported in the AWUPF field of the Identify Controller data structure. All other values specify a size in terms of logical blocks using the same encoding as the AWUPF field.	0x00

continued...

Identify Namespace Data Structure, NVM Command Set Specific			
Bytes	**O/M**	**Description**	**Expected Value**
39:38	O	Namespace Atomic Compare & Write Unit (NACWU): This field indicates the namespace specific size of the write operation guaranteed to be written atomically to the NVM for a Compare and Write fused command. A value of 0h indicates that the size for this namespace is the same size as that reported in the ACWU field of the Identify Controller data structure. All other values specify a size in terms of logical blocks using the same encoding as the ACWU field.	0x00
41:40	O	Namespace Atomic Boundary Size Normal (NABSN): This field indicates the atomic boundary size for this namespace for the NAWUN value. This field is specified in logical blocks. Writes to this namespace that cross atomic boundaries are not guaranteed to be atomic to the NVM with respect to other read or write commands. A value of 0h indicates that there are no atomic boundaries for normal write operations. All other values specify a size in terms of logical blocks using the same encoding as the AWUN field.	0x00
43:42	O	Namespace Atomic Boundary Offset (NABO): This field indicates the LBA on this namespace where the first atomic boundary starts. If the NABSN and NABSPF fields are cleared to 0h, then the NABO field shall be cleared to 0h. NABO shall be less than or equal to NABSN and NABSPF.	0x00
45:44	O	Namespace Atomic Boundary Size Power Fail (NABSPF): This field indicates the atomic boundary size for this namespace specific to the Namespace Atomic Write Unit Power Fail value. This field is specified in logical blocks. Writes to this namespace that cross atomic boundaries are not guaranteed to be atomic with respect to other read or write commands and there is no guarantee of data returned on subsequent reads of the associated logical blocks. A value of 0h indicates that there are no atomic boundaries for power fail or error conditions. All other values specify a size in terms of logical blocks using the same encoding as the AWUPF field.	0x00
47:46		Reserved	
63:48	O	NVM Capacity (NVMCAP): This field indicates the total size of the NVM allocated to this namespace. The value is in bytes. This field shall be supported if Namespace Management and Namespace Attachment commands are supported. ***Note:*** This field may not correspond to the logical block size multiplied by the Namespace Size field. Due to thin provisioning or other settings (e.g., endurance), this field may be larger or smaller than the Namespace Size reported.	Varies
103:64		Reserved	
119:104	O	Namespace Globally Unique Identifier (NGUID): This field contains a 128-bit value that is globally unique and assigned to the namespace when the namespace is created. This field remains fixed throughout the life of the namespace and is preserved across namespace and controller operations (e.g., controller reset, namespace format, etc.). This field uses the EUI-64 based 16-byte designator format. Bytes 114:112 contain the 24- bit Organizationally Unique Identifier (OUI) value assigned by the IEEE Registration Authority. Bytes 119:115 contain an extension identifier assigned by the corresponding organization. Bytes 111:104 contain the vendor specific extension identifier assigned by the corresponding organization. See the IEEE EUI-64 guidelines for more information. The controller shall specify a globally unique namespace identifier in this field or the EUI64 field when the namespace is created.	Varies based on WWN & NSID
127:120		IEEE Extended Unique Identifier (EUI64): This field contains a 64-bit IEEE Extended Unique Identifier (EUI-64) that is globally unique and assigned to the namespace when the namespace is created. This field remains fixed throughout the life of the namespace and is preserved across namespace and controller operations (e.g., controller reset, namespace format, etc.).	0x00

continued...

Identify Namespace Data Structure, NVM Command Set Specific			
Bytes	**O/M**	**Description**	**Expected Value**
		The EUI-64 is a concatenation of a 24-bit or 36-bit Organizationally Unique Identifier (OUI or OUI-36) value assigned by the IEEE Registration Authority and an extension identifier assigned by the corresponding organization. See the IEEE EUI-64 guidelines for more information. The controller shall specify a globally unique namespace identifier in this field or the NGUID field when the namespace is created. If the controller is not able to allocate a globally unique 64-bit identifier then this field shall be cleared to 0h.	
131:128	M	LBA Format 0 Support (LBAF0): This field indicates the LBA format 0 that is supported by the controller.	LBA Format Data Structure in Table 47 on page 56
135:132	O	LBA Format 1 Support (LBAF1): This field indicates the LBA format 1 that is supported by the controller.	
139:136	O	LBA Format 2 Support (LBAF2): This field indicates the LBA format 2 that is supported by the controller.	
143:140	O	LBA Format 3 Support (LBAF3): This field indicates the LBA format 3 that is supported by the controller.	
147:144	O	LBA Format 4 Support (LBAF4): This field indicates the LBA format 4 that is supported by the controller.	
151:148	O	LBA Format 5 Support (LBAF5): This field indicates the LBA format 5 that is supported by the controller.	
155:152	O	LBA Format 6 Support (LBAF6): This field indicates the LBA format 6 that is supported by the controller.	
159:156	O	LBA Format 7 Support (LBAF7): This field indicates the LBA format 7 that is supported by the controller.	
163:160	O	LBA Format 8 Support (LBAF8): This field indicates the LBA format 8 that is supported by the controller.	
167:164	O	LBA Format 9 Support (LBAF9): This field indicates the LBA format 9 that is supported by the controller.	
171:168	O	LBA Format 10 Support (LBAF10): This field indicates the LBA format 10 that is supported by the controller.	
175:172	O	LBA Format 11 Support (LBAF11): This field indicates the LBA format 11 that is supported by the controller.	
179:176	O	LBA Format 12 Support (LBAF12): This field indicates the LBA format 12 that is supported by the controller.	
183:180	O	LBA Format 13 Support (LBAF13): This field indicates the LBA format 13 that is supported by the controller.	
187:184	O	LBA Format 14 Support (LBAF14): This field indicates the LBA format 14 that is supported by the controller.	
191:188	O	LBA Format 15 Support (LBAF15): This field indicates the LBA format 15 that is supported by the controller.	
383:192		Reserved	
4095:384	O	Vendor Specific (VS): This range of bytes is allocated for vendor specific usage.	

Note: O = Optional. The content of the word is optional

M = Mandatory. The content of the word is mandatory

Table 47: LBA Format Data Structure – 512B

Bits	Description	512B	4096B
31:26	Reserved		
25:24	**Relative Performance (RP):** This field indicates the relative performance of the LBA format indicated relative to other LBA formats supported by the controller. Depending on the size of the LBA and associated metadata, there may be performance implications. The performance analysis is based on better performance on a queue depth 32 with 4KB read workload. The meanings of the values indicated are included in the following table.	0x02	0x00
	Value / **Definition** 00b – Best performance 01b – Better performance 10b – Good performance 11b – Degraded performance		
23:16	**LBA Data Size (LBADS):** This field indicates the LBA data size supported. The value is reported in terms of a power of two (2^n). A value smaller than 9 (i.e. 512 bytes) is not supported. ***Note:*** If the value reported is 0h then the LBA format is not supported / used.	0x09	0x0C
15:00	**Metadata Size (MS):** This field indicates the number of metadata bytes provided per LBA based on the LBA Size indicated. The namespace may support the metadata being transferred as part of an extended data LBA or as part of a separate contiguous buffer. If end-to-end data protection is enabled, then the first eight bytes or last eight bytes of the metadata is the protection information.	0x00	0x00

Table 48: Possible DPC & DPS values

Bits	Description	512B	4096B
28	**End-to-end Data Protection Capabilities (DPC):** This field indicates the capabilities for the end-to-end data protection feature. Multiple bits may be set in this field.	x00	x00
	Bits 7:5 are reserved		
	Bit 4 if set to '1' indicates that the namespace supports protection information transferred as the last eight bytes of metadata. Bit 4 if cleared to '0' indicates that the namespace does not support protection information transferred as the last eight bytes of metadata	0	0
	Bit 3 if set to '1' indicates that the namespace supports protection information transferred as the first eight bytes of metadata. Bit 3 if cleared to '0' indicates that the namespace does not support protection information transferred as the first eight bytes of metadata.	0	0
	Bit 2 if set to '1' indicates that the namespace supports Protection Information Type 3. Bit 2 if cleared to '0' indicates that the namespace does not support Protection Information Type 3.	0	0
	Bit 1 if set to '1' indicates that the namespace supports Protection Information Type 2. Bit 1 if cleared to '0' indicates that the namespace does not support Protection Information Type 2.	0	0
	Bit 0 if set to '1' indicates that the namespace supports Protection Information Type 1. Bit 0 if cleared to '0' indicates that the namespace does not support Protection Information Type 1.	0	0

§

D Vital Data Structure (0x53)

Each NVM Subsystem with one or more Management Endpoints shall have a FRU Information Device which is compliant with the IPMI Platform Management FRU Information Storage Definition. The VPD shall contain the required elements defined in table below. The size of the VPD is 256 bytes as defined by the IPMI Platform Management FRU Information Storage Definition.

Refer to Appendix A of the NVMe-MI 1.0 Specification on www.nvme.org for Basic Command Management description.

Table 49: Vital Product Data Structure (VPD Elements)

Byte	Name
07:00	Common Header
119:08	Product Info Area
Vendor Specific:120	MultiRecord Info Area
Vendor Specific	Internal Use Area (optional)
Vendor Specific	Chassis Info Area (optional)
Vendor Specific	Board Info Area (optional)

The VPD shall be accessible using the VPD Read command. The entire contents of the VPD may be updated using the VPD Write command.

If the NVM Subsystem has a SMBus/I2C interface, the VPD shall be accessible at the SMBus/I2C address of the FRU Information Device using the access mechanism over I2C as defined in the IPMI Platform Management FRU Information Storage Definition. Updating the VPD by writing to the FRU Information Device directly on SMBus/I2C shall not be supported.

VPD records utilize the Type/Length byte format defined in the IPMI Platform Management FRU Information Storage Definition. Type/Length byte encodings utilized in this specification are summarized in the following table

Table 50: Type/Length Byte Format

Bits	Field Name	Description
7:6	Type Code	Specifies field encoding 11b – Always corresponds to ASCII in this specification
5:0	Number of Data Bytes	Specifies field length 000000b indicates that th field is empty

Table 51: Common Header

Byte	Factory Default	Description
0	01h	**IPMI Format Version Number (IPMIVER)**: This field indicates the IPMI Format Version.
1	00h	**Internal Use Area Starting Offset (IUAOFF):** This field indicates the starting offset in multiples of 8 bytes for the Internal Use Area. A value of 00h may be used to indicate the Internal Use Area is not present.
2	00h	**Chassis Info Area Starting Offset (CIAOFF)**: This field indicates the starting offset in multiples of 8 bytes for the Chassis Info Area. A value of 00h may be used to indicate the Chassis Info Area is not present.

continued...

Byte	Factory Default	Description
3	00h	**Board Info Area Starting Offset (BIAOFF)**: This field indicates the starting offset in multiples of 8 bytes for the Board Info Area. A value of 00h may be used to indicate the Board Info Area is not present.
4	01h	**Product Info Area Starting Offset (PIAOFF)**: This field indicates the starting offset in multiples of 8 bytes for the Product Info Area.
5	0Fh	**MultiRecord Info Area Starting Offset (MRIOFF):** This field indicates the starting offset in multiples of 8 bytes for the MultiRecord Info Area.
6	00h	Reserved
7	Impl Spec	**Common Header Checksum (CHCHK):** Checksum computed over bytes 0 through 6. The checksum is computed by adding the 8-bit value of the bytes modulo 256 and then taking the 2's complement of this sum. When the checksum and the sum of the bytes module 256 are added, the result should be 0h.

Table 52: Product Info Area (offset 8 bytes)

Byte Offset	Factory Default	Description
0	01h	**IPMI Format Version Number (IPMIVER)**: This field indicates the IPMI Format Version.
1	0Eh	**Product Info Area Length (PALEN)**: This field indicates the length of the product info area in multiples of 8 bytes. 112 bytes/8 = 14 = 0x0Eh
2	19h	**Language Code (LCODE):** This field indicates the language used. A value of 19h is used to indicate English.
3	C8h	**Manufacturer Name Type/Length (MNTL):** This byte indicates the type and length of the Manufacturer Name field.
11:04	0x49 0x4E 0x54 0x45 0x4C 0x00 0x00 0x00	**Manufacturer Name (MNAME):** This field indicates the Manufacturer name in 8-bit ASCII. Unused bytes should be NULL characters. The MNAME value shall correspond to "INTEL" in ASCII padded with 0x00.
12	D8h	**Product Name Type/Length (PNTL):** This byte indicates the type and length of the Product Name field.
36:13	Impl Spec	**Product Name (PNAME):** This field indicates the Product name in 8-bit ASCII. Unused bytes should be NULL characters.
37	E8h	**Product Part/Model Number Type/Length (PPMNNTL):** This byte indicates the type and length of the Product Part/Model Number field.
77:38	Varies depending on SKU	**Product Part/Model Number (PPMN):** This field indicates the Product Part/Model Number in 8-bit ASCII. Unused bytes should be NULL characters.
78	C2h	**Product Version Type/Length (PVTL):** This byte indicates the type and length of the Product Part/Model Number field.
80:79	D4h	**Product Version (PVER):** This field indicates the Product Version in 8-bit ASCII. Unused bytes should be NULL characters.
81	D4h	**Product Serial Number Type/Length (PSNTL):** This byte indicates the type and length of the Product Serial Number field.
101:82	Impl Spec	This field should contain the same value as the Model Number (NM) field in the NVMe Identify Controller Data Structure
102	0h	**Asset Tag Type/Length (ATTL):** This byte indicates the type and length of the Asset Tag field. A value of 00h may be used to indicate an Asset Tag is not present.

continued...

Byte Offset	Factory Default	Description
103	0h	**FRU File ID Type/Length (ATTL):** This byte indicates the type and length of the FRU File ID field. A value of 00h may be used to indicate a FRU File ID is not present.
104	C1h	**End of Record (EOR)**: A value of C1h in this byte indicates the end of record
110:105		Reserved
111	Impl Spec	**Product Info Area (PICHK):** Checksum computed over bytes 0 through 110. The checksum is computed by adding the 8-bit value of the bytes modulo 256 and then taking the 2's complement of this sum. When the checksum and the sum of the bytes module 256 are added, the result should be 0h.

Table 53: NVMe MultiRecord Area

Byte Offset	Factory Default	Description
0	0Bh	NVMe Record Type
1	02h	Bit 7 – end of list; record format version = 2h
2	28h	**Record Length (RLEN)**: This field indicates the length of the MultiRecord Area in bytes.
3	Impl Spec	**Record Checksum**: This field is used to give the record data a zero checksum (i.e., the modulo 256 sum of the record data bytes from byte offset 05 through the end of this record plus this checksum byte equals zero)
4	Impl Spec	**Header Checksum:** This field is used to give the record header a zero checksum (i.e., the modulo 256 sum of the preceding record bytes starting with the first byte of the header plus this checksum byte equals zero.
5	0h	**NVMe MultiRecord Area Version Number:** This field indicates the version number of this multirecord. This field shall be set to 0h in this version of the specification.
06	Impl Spec	**Management Endpoint Form Factor (MEFF):** This field indicates the form factor of the Management Endpoint.

Value	Definition
0	Other – unknown
1 – 15	Reserved
16	2.5" Form Factor – unknown
17	2.5" Form Factor – U.2 (SFF-8639) 15mm
18	2.5" Form Factor – U.2 (SFF-8639) 7mm
19 – 31	Reserved
32	CEM add in card – unknown
33	CEM add in card – Low Profile (HHHL)
34	CEM add in card – Standard Height Half Length (FHHL)
35	CEM add in card – Standard Height Full Length (FHFL)
36-47	Reserved
48	M.2 module – unknown
49	M.2 module – 2230
50	M.2 module – 2242
51	M.2 module – 2260
52	M.2 module – 2280

continued...

Byte Offset	Factory Default	Description
		<table><tr><th>Value</th><th>Definition</th></tr><tr><td>53</td><td>M.2 module – 22110</td></tr><tr><td>54-63</td><td>Reserved</td></tr><tr><td>64</td><td>BGA SSD – unknown</td></tr><tr><td>65</td><td>BGA SSD – 16 x 20mm (M.2 Type 1620)</td></tr><tr><td>66</td><td>BGA SSD – 16 x 20mm (M.2 Type 1113)</td></tr><tr><td>67-79</td><td>Reserved</td></tr><tr><td>80</td><td>Enterprise and Datacenter SSD Form Factor – unknown</td></tr><tr><td>81</td><td>E1.S Form Factor – (SFF-TA-1006) 5.9mm</td></tr><tr><td>82</td><td>E1.S Form Factor – (SFF-TA-1006) 8mm</td></tr><tr><td>83</td><td>E1.L Form Factor – (SFF-TA-1007) 9.5mm</td></tr><tr><td>84</td><td>E1.L Form Factor – (SFF-TA-1007) 18mm</td></tr><tr><td>85</td><td>3" Short Form Factor – (SFF-TA-1008) 7.5mm</td></tr><tr><td>86</td><td>3" Short Form Factor – (SFF-TA-1008) 16.8mm</td></tr><tr><td>87</td><td>3" Long Form Factor – (SFF-TA-1008) 7.5mm</td></tr><tr><td>88</td><td>3" Long Form Factor – (SFF-TA-1008) 16.8mm</td></tr><tr><td>89-239</td><td>Reserved</td></tr><tr><td>240-255</td><td>Vendor Specific</td></tr></table>
12:07		Reserved
13	00h	**Initial 1.8V Power Supply Requirements:** This field specifies the initial 1.8V power supply requirements in Watts prior to receiving a Set Slot Power message.
14	Impl Spec	**Maximum 1.8V Power Supply Requirements:** This field specifies the maximum 1.8V power supply requirements in Watts. A value of zero indicates that the power supply voltage is not used.
15	Impl Spec	**Initial 3.3V Power Supply Requirements:** This field specifies the initial 3.3V power supply requirements in Watts prior to receiving a Set Slot Power message. 00h – U.2, Add-in card 08h – M.2
16	Impl Spec	**Maximum 3.3V Power Supply Requirements:** This field specifies the maximum 3.3V power supply requirements in Watts. A value of zero indicates that the power supply voltage is not used.
17	Impl Spec	Reserved
18	Impl Spec	**Maximum 3.3Vaux Power Supply Requirements:** This field specifies the maximum 3.3V power supply requirements in 10 mW units. A value of zero indicates that the power supply voltage is not used.
19	Impl Spec	**Initial 5V Power Supply Requirements:** This field specifies the initial 5V power supply requirements in Watts prior to receiving a Set Slot Power message.
20	Impl Spec	**Maximum 5V Power Supply Requirements:** This field specifies the maximum 5V power supply requirements in Watts. A value of zero indicates that the power supply voltage is not used.
21	Impl Spec	**Initial 12V Power Supply Requirements:** This field specifies the initial 12V power supply requirements in Watts prior to receiving a Set Slot Power message.

continued...

Byte Offset	Factory Default	Description
22	Impl Spec	**Maximum 12V Power Supply Requirements:** This field specifies the maximum 12V power supply requirements in Watts. A value of zero indicates that the power supply voltage is not used.
23	Impl Spec	**Maximum Thermal Load:** This field specifies the maximum thermal load from the NVM Subsystem in Watts.
36:24	Same as TNVMCAP	**Total NVM Capacity:** This field indicates the total NVM capcity of the Management Endpoint in bytes. If the NVM Subsystem supports Namespace Management, then this field should correspond to the value reported in the TNVMCAP field in the NVMe Identify Controller Data structure. A value of 0h may be used to indicate this feature is not supported.
63:37		Reserved

Table 54: NVMe PCIe Port MultiRecord Area

Byte Offset	Factory Default	Description
0	0Ch	**NVMe PCIe Port Record Type ID**
1	2h	Bit 7 – end of list; record format version = 2h
2	28h	**Record Length (RLEN)**: This field indicates the length of the MultiRecord Area in bytes.
3	Impl spec	**Record Checksum: This field is used to give the record data a zero checksum (i.e., the modulo 256 sum of the record data bytes from byte offset 05 through the end of this record plus this checksum byte equals zero)**
04	Spec	**Header Checksum**: This field is used to give the record header a zero checksum (i.e., the modulo 256 sum of the preceding record bytes starting with the first byte of the header plus this checksum byte equals zero.
5	0h	**NVMe PCIe Port MultiRecord Area Version Number:** This field indicates the version number of this multirecord. This field shall be set to zero in this version of the specification.
6	(PN) field of PXLCAP	**PCIe Port Number:** This field contains the PCIe port number. This is the same value as that reported in the Port Number field in the PCIe Link Capabilities Register.
7	01h	**Port Information:** This field indicates information about the PCIe Ports in the device. Bits 7 to 1 are reserved. Bit 0, if set to '1' indicates that all PCIe ports within the device have the same capabilities (i.e., the capabilities listed in this structure are consistent across each PCIe port). Single port.
08	07h	**PCIe Link Speed:** This field indicates a bit vector of link speeds supported by the PCIe port. Bit 7:3 – Reserved Bit 2 – Set to '1' if the PCIe link supports 8.0 GT/s. Otherwise cleared to '0'. Bit 1 – Set to '1' if the PCIe link supports 5.0 GT/s. Otherwise cleared to '0'. Bit 0 – Set to '1' if the PCIe link supports 2.5 GT/s. Otherwise cleared to '0'.
09	0x04h	**PCIe Maximum Link Width:** The maximum PCIe link width for this NVM Subsystem port. This is the expected negotiated link width that the port link trains to if the platform supports it. A Management Controller may compare this value with the PCIe Negotiated Link Width to determine if there has been a PCIe link training issue. Value 0 – Reserved Value 1 – PCIe x1 Value 2 – PCIe x2

continued...

<table>
<tr><th>Byte Offset</th><th>Factory Default</th><th>Description</th></tr>
<tr><td></td><td></td><td>
<table>
<tr><th>Value</th><th>Definition</th></tr>
<tr><td>3</td><td>Reserved</td></tr>
<tr><td>4</td><td>PCIe x4</td></tr>
<tr><td>5-7</td><td>Reserved</td></tr>
<tr><td>8</td><td>PCIe x8</td></tr>
<tr><td>9-11</td><td>Reserved</td></tr>
<tr><td>12</td><td>PCIe x12</td></tr>
<tr><td>13-15</td><td>Reserved</td></tr>
<tr><td>16</td><td>PCIe x16</td></tr>
<tr><td>17-31</td><td>Reserved</td></tr>
<tr><td>32</td><td>PCIe x32</td></tr>
<tr><td>33-255</td><td>Reserved</td></tr>
</table>
</td></tr>
<tr><td>10</td><td>0h</td><td>MCTP Support: This field contains a bit vector that specifies the level of support for the NVMe Management Interface.
Bits 7 to 1 are reserved.
Bit 0, if set to ‘1’ indicates that MCTP based management commands are supported on the PCIe port.</td></tr>
<tr><td>11</td><td>Impl Spec</td><td>Ref Clk Capability: This field contains a bit vector that specifies the PCIe clocking modes supported by the port.
<table>
<tr><th>Bit</th><th>Definition</th></tr>
<tr><td>7:4</td><td>Reserved</td></tr>
<tr><td>3</td><td>Set to ‘1’ if the device automatically uses RefClk if provided and otherwise uses SRIS. Otherwise cleared to ‘0’.</td></tr>
<tr><td>2</td><td>Set to ‘1’ if the PCIe link supports Separate ReClk with SSC (SRIS). Otherwise cleared to ‘0’.</td></tr>
<tr><td>1</td><td>Set to ‘1’ if the PCIe link supports Separate ReClk with no SSC (SRNS). Otherwise cleared to ‘0’.</td></tr>
<tr><td>0</td><td>Set to ‘1’ if the PCIe link supports common ReClk. Otherwise cleared to ‘0’.</td></tr>
</table>
</td></tr>
<tr><td>15:12</td><td>00h</td><td>Reserved</td></tr>
</table>

§

E Out-of-Band Command Response Using SMBus (0x6A)

The SMBus slave address to read this data structure is allowed to be the same address we use for MCTP, and defaults to 0x6Ah

Refer to Appendix A of the NVMe-MI 1.0 Specification on www.nvme.org for Basic Command Management description.

Table 55: System Management Data Structure (NVMe-MI Commands)

<table>
<tr><th>Command Code</th><th>Offset (byte)</th><th>Description</th></tr>
<tr><td rowspan="4">0</td><td>00</td><td>Length of Status: Indicates number of additional bytes to read before encountering PEC. This value should always be 6 (06h) in implementations of this version of the spec.</td></tr>
<tr><td>01</td><td>Status Flags (SFLGS): This field indicates the status of the NVM subsystem.
SMBus Arbitration – Bit 7 is set ‘1’ after a SMBus block read is completed all the way to the stop bit without bus contention and cleared to ‘0’ if a SMBus Send Byte FFh is received on this SMBus slave address.
Drive Not Ready – Bit 6 is set to ‘1’ when the subsystem cannot process NVMe management commands, and the rest of the transmission may be invalid. If cleared to ‘0’, then the NVM subsystem is fully powered and ready to respond to management commands. This logic level intentionally identifies and prioritizes powered up and ready drives over their powered off neighbors on the same SMBus segment.
Drive Functional – Bit 5 is set to ‘1’ to indicate an NVM subsystem is functional. If cleared to ‘0’, then there is an unrecoverable failure in the NVM subsystem and the rest of the transmission may be invalid.
Reset Not Required - Bit 4 is set to ‘1’ to indicate the NVM subsystem does not need a reset to resume normal operation. If cleared to ‘0’ then the NVM subsystem has experienced an error that prevents continued normal operation. A controller reset is required to resume normal operation.
Port 0 PCIe Link Active - Bit 3 is set to ‘1’ to indicate the first port’s PCIe link is up (i.e., the Data Link Control and Management State Machine is in the DL_Active state). If cleared to ‘0’, then the PCIe link is down.
Port 1 PCIe Link Active - Bit 2 is set to ‘1’ to indicate the second port’s PCIe link is up. If cleared to ‘0’, then the second port’s PCIe link is down or not present.
Bits 1-0 shall be set to ‘1’.</td></tr>
<tr><td>02</td><td>SMART Warnings: This field shall contain the Critical Warning field (byte 0) of the NVMe SMART / Health Information log. Each bit in this field shall be inverted from the NVMe definition (i.e., the management interface shall indicate a ‘0’ value while the corresponding bit is ‘1’ in the log page). Refer to the NVMe specification for bit definitions.
If there are multiple controllers in the NVM subsystem, the management endpoint shall combine the Critical Warning field from every controller such that a bit in this field is:
• Cleared to ‘0’ if any controller in the subsystem indicates a critical warning for that corresponding bit.
Set to ‘1’ if all controllers in the NVM subsystem do not indicate a critical warning for the corresponding bit.</td></tr>
<tr><td>03</td><td>Composite Temperature (CTemp): This field indicates the current temperature in degrees Celsius. If a temperature value is reported, it should be the same temperature as the Composite Temperature from the SMART log of hottest controller in the NVM subsystem. The reported temperature range is vendor specific, and shall not exceed the range -60 to +127° C. The 8 bit format of the data is shown below.
This field should not report a temperature when that is older than 5 seconds. If recent data is not available, the NVMe management endpoint should indicate a value of 80h for this field.
<table>
<tr><th>Value</th><th>Description</th></tr>
<tr><td>00h-7Eh</td><td>Temperature is measured in degrees Celsius (0° to 126° C)</td></tr>
<tr><td>7Fh</td><td>127° C or higher</td></tr>
<tr><td>80h</td><td>No temperature data or temperature data is more than 5 seconds old.</td></tr>
</table></td></tr>
</table>

continued...

<table>
<tr><th>Command Code</th><th>Offset (byte)</th><th>Description</th></tr>
<tr><td rowspan="5"></td><td rowspan="2"></td><td><table>
<tr><th>Value</th><th>Description</th></tr>
<tr><td>81h</td><td>Temperature sensor failure</td></tr>
<tr><td>82h-C3h</td><td>Reserved</td></tr>
<tr><td>C4</td><td>Temperature is -60° C or lower</td></tr>
<tr><td>C5-FFh</td><td>Temperature measured in degrees Celsius is represented in twos complement (-1° to -59° C)</td></tr>
</table></td></tr>
<tr><td>04</td><td>Percentage Drive Life Used (PDLU): Contains a vendor specific estimate of the percentage of NVM subsystem NVM life used based on the actual usage and the manufacturer's prediction of NVM life. If an NVM subsystem has multiple controllers the highest value is returned. A value of 100 indicates that the estimated endurance of the NVM in the NVM subsystem has been consumed, but may not indicate an NVM subsystem failure. The value is allowed to exceed 100. Percentages greater than 254 shall be represented as 255. This value should be updated once per power-on hour and equal the Percentage Used value in the NVMe SMART Health Log Page.</td></tr>
<tr><td>06:05</td><td>Reserved: Shall be set to 0000h</td></tr>
<tr><td>07</td><td>PEC: An 8 bit CRC calculated over the slave address, command code, second slave address and returned data. Algorithm is in SMBus Specifications.</td></tr>
<tr><td rowspan="4">8</td><td>08</td><td>Length of identification: Indicates number of additional bytes to read before encountering PEC. This value should always be 22 (16h) in implementations of this version of the spec.</td></tr>
<tr><td>10:09</td><td>Vendor ID: The 2 byte vendor ID, assigned by the PCI SIG. Should match VID in the Identify Controller command response. MSB is transmitted first.</td></tr>
<tr><td>30:11</td><td>Serial Number: 20 characters that match the serial number in the NVMe Identify Controller command response. First character is transmitted first</td></tr>
<tr><td>31</td><td>PEC: An 8 bit CRC calculated over the slave address, command code, second slave address and returned data. Algorithm is in SMBus Specifications.</td></tr>
<tr><td>32+</td><td>255:32</td><td>Vendor Specific – This data structure shall not exceed the maximum read length of 255 specified in the SMBus version 3 specification. Preferably length is not greater than 32 for compatibility with SMBus 2.0, additional blocks shall be on 8 byte boundaries.</td></tr>
</table>

§

F Out-of-Band Command Response Using SMBus (0x6A Intel Specific)

Table 56: Command Response 0x6A (Intel Specific Vendor Unique Commands)

Command Code	Offset (byte)	Description
32	32	**Length of Intel Corporation's Block:** shall be 22 until this spec is updated
	33	**Reserved:** Currently cleared to zero. Anticipated future use will be several bits for version number, a bit if there are additional Intel blocks, and some other less thought out ideas at this point
	34	**Reserved:** Value = 0xff
	35	**Reserved:** Value = 0x80
	38:36	**Reserved:** Value = 0x000000
	46:39	**Firmware Version:** 8 characters, ASCII representation
	54:47	**Bootloader Version:** 8 characters, ASCII representation
	55	**PEC:** An 8 bit CRC calculated over the slave address, command code, second slave address and returned data. Algorithm is in SMBus Specifications.
56+	255:56	**Reserved:** all bytes cleared to zero, no PEC

§

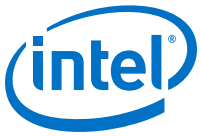
G SCSI Command Translation

Following SCSI commands are supported:

- Read 6,10,12,16
- Inquiry
- Mode Sense 6,10
- Mode Select 6.10
- Log Sense
- Read Capacity 10,16
- Report LUNs
- Request Sense
- Start Stop Unit
- Test Unit Ready
- Write Buffer
- Unmap

Note: Refer to NVM Express: SCSI translation reference doc under nvmexpress.org

§



H PCIe* ID

Table 57: PCIe ID

ID Name	Description	U.2 15mm	PCIe Register Location	Identify Controller Location	Vital Product Data Location
Vendor ID (VID)	Vendor ID assigned by PCI-SIG	0x8086	PCI Header Offset 00h (bits 15:00)	Bytes 01:00h	Address 3, (size 2B)
Device ID (DID)	Device ID assigned by vendor	0x0A54	PCI Header Offset 00h (bits 31:16)	NA	NA
Subsystem Vendor ID	Indicates Sub-system vendor ID	0x8086	PCI Header Offset 2Ch (bits 15:00)	Bytes 03:02h	NA
Subsystem ID	Sub-system identifier	0x4308	PCI Header Offset 2Ch (bits 31:16)	NA	NA

§