# Intel® Solid-State Drive DC P3608 Series

## Product Specification

- Capacities:
  - 1.6TB, 3.2TB, 4.0TB
- Components
  - Intel® 20nm [HET High Endurance Technology] MLC NAND Flash Memory
- PCIe® 3.0, x8
- Form Factors
  - AIC Form Factor (Add-in Card)
    - Half-height, Half-length, Low Profile
- Performance[1,2]
  - Seq R/W: Up to 5000/3000MB/s[3]
  - IOPS Random 70/30 R/W 4KB[4]: Up to 300K
  - IOPS Random 4KB[4] R/W: Up to 850/150K
  - Seq Latency (typ) R/W: 20/20µs
- Operating System Support:
  - Windows* Server 2012 R2, 2012, 2008 R2 x64
  - Windows 7*, Windows 8*, Windows 8.1* (32bit/64bit)
  - RHEL* 6.5, 6.6, 6.7, 7.0, 7.1
  - UEFI 2.3.1*
  - SLES11SP4*, SLES12
  - Citrix Xenserver 6.5, Cent OS 6.7
  - VMWare ESXI.5.5
- Reliability
  - Uncorrectable Bit Error Rate (UBER): 1E-17
  - Silent bit error rate of 1E-25
  - Mean Time Before Failure (MTBF): 1 million hours
  - T10 DIF protection
- Variable Sector Size:
  - 512, 520, 528, 4096, 4104, 4160, 4224 Bytes
- Power
  - 3.3V and 12V Supply Rail
  - Enhanced power-loss data protection
  - Active (TYP): 40W
  - Idle (TYP): 11.5W
- Compliance
  - NVM Express™ 1.0
  - PCI Express® Base Specification Rev 3.0
  - Enterprise SSD Form Factor Version 1.0a
  - PCI Express Card Electro-Mechanical (CEM) Specification Rev 3.0
  - Microsoft Windows Hardware Certification*
- Certifications and Declarations
  - UL*, CE*, C-Tick*, BSMI*, KCC*, VCCI*
- Endurance Rating
  - Up to 21.9 PBW (Petabytes Written)[5] 3 Drive Writes/day (JESD219 workload)
- Temperature Specification
  - Operating:: 0 to 55° C with specified airflow
  - Non-Operating[6]: -55 to 95° C
  - Temperature monitoring (In-band through NVMe SMART and out of band)
  - Thermal throttling when approaching maximum operating temperature
- Airflow
  - 55° C @ 400LFM (for 40W drive airflow towards IO bracket[7])
- Weight
  - 1.6TB 227.5g, 3.2TB 230.2g, 4.0TB 230.6g
- Shock
  - 50 G Trapezoidal, 170 in/s
- Vibration
  - 3.13 $G_{RMS}$ (5-500Hz)
- Altitude (Simulated)
  - Operating: -1,000 to 10,000 ft
  - Non-Operating: -1,000 to 40,000 ft
- Product Ecological Compliance
  - RoHS

1. Performance values vary by capacity and form factor and are an aggregate value of the two volumes
2. Performance specifications apply to both compressible and incompressible data
3. MB/s = 1,000,000 bytes/second
4. 4KB = 4,096 bytes; 8 KB = 8,192 bytes
5. 1PB = $10^{15}$ Bytes; per volume.
6. Please contact your Intel representative for details on the non-operating temperature range
7. Airflow out of server through PCIe Card Slot

Order Number: 333055-001

# Ordering Information

Contact your local Intel sales representative for ordering information.

# Revision History

| Revision Number | Description | Revision Date |
|---|---|---|
| 001 | • Initial release | September 2015 |

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.

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. For more complete information about performance and benchmark results, visit http://www.intel.com/performance

**Test System Configuration:** Intel® Core™ i7-4770 CPU at 3.4GHz, 8 GB DDR3 at 1600MHz, Intel® SSD DC P3608 Series 1.6TB

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at www.intel.com.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

All documented test results are obtained in compliance with JESD218 Standards; refer to individual sub-sections within this document for specific methodologies. See www.jedec.org for detailed definitions of JESD218 Standards.

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.

Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to: http://www.intel.com/design/literature.htmIntel and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

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

# Contents


1 Overview ..... 5
1.1 References ..... 6
1.2 Terms and Acronyms ..... 7
2 Product Specifications ..... 8
2.1 Capacity ..... 8
2.2 Performance ..... 8
2.3 Electrical Characteristics ..... 10
2.4 Environmental Conditions ..... 12
2.5 Product Regulatory Compliance ..... 13
2.6 Reliability Specifications ..... 14
2.7 Temperature Sensor ..... 14
2.8 Power Loss Capacitor Test ..... 15
2.10 Out of Band Management (SMBUS) ..... 15
2.11 Variable Sector Size Support ..... 15
3 Mechanical Information ..... 16
4 Pin and Signal Descriptions ..... 17
4.1 AIC Pin Definition Table ..... 17
5 Supported Command Sets ..... 19
5.1 NVMe Admin Command Set ..... 19
5.2 NVMe I/O Command Set ..... 19
5.3 Log Page Support ..... 20
5.4 SMART Attributes ..... 20
5.5 Temperature Statistics ..... 24
5.6 Drive Marketing Name Log ..... 24
5.7 IO Command Latency Statistics ..... 25
5.7 SET Feature Identifiers ..... 25
6 NVMe Driver Support ..... 28
7 Certifications and Declarations ..... 29
Appendix A IDENTIFY Data Structure ..... 30
Appendix B Vital Data Structure ..... 38
Appendix C Out of Band Temperature Sensor Read Out ..... 39
Appendix D SMBUS Command Response on 0x6A / 0x6B (NVMe Workgroup Defined) ..... 40
Appendix E SMBUS Command Response on 0x6A and 0x6B (Intel Specific) ..... 43
Appendix F PCIe ID ..... 44
Appendix G SCSI Command Translation ..... 45
Appendix H Add-in Card LED Decoder ..... 46

## Figures

Figure 3-1: Intel SSD DC P3608 Series Dimensions ..... 16
Figure F-1: LED Location ..... 46

## Tables

Table 1: Standard Information Referenced in this Document ..... 6
Table 2: Glossary of Terms and Acronyms ..... 7
Table 3: User Addressable Sectors ..... 8
Table 4: Random Read/Write Input/Output Operations Per Second (IOPS) ..... 8
Table 5: Random Read/Write IOPS Consistency ..... 9
Table 6: Sequential Read and Write Bandwidth ..... 9
Table 7: Latency ..... 9
Table 8: Quality of Service ..... 10
Table 9: Operating Voltage ..... 10
Table 10: Power Consumption ..... 11
Table 11: Temperature, Shock, Vibration ..... 12
Table 12: Airflow Requirements for Intel SSD DC P3608 Series (Add-In Card) ..... 12
Table 13: Product Regulatory Compliance Specifications ..... 13
Table 14: Reliability Specifications ..... 14
Table 15: Pin Definition for Add-In Card (Half Height Half Length) Form Factor ..... 17
Table 16: SMART Attributes (Log Identifier 02h) ..... 20
Table 17: Additional SMART Attributes (Log Identifier CAh) ..... 22
Table 18: Temperature Statistics (Log Identifier C5h) ..... 24
Table 19: Drive Marketing Name Log (Log Identifier DDh) ..... 24
Table 20: Read/Write Command Latency Log (Log Identifier C1h/C2h) ..... 25
Table 21: Set Max LBA Setting - Command Dword 11 and Command Dword 12 ..... 25
Table 22: Status Code - Set Max LBA Command Specific Status Values ..... 26
Table 23: C6h - Set/ Get Power (Typical) Governor Setting – Command Dword 11 ..... 26
Table 24: Status Codes - Power Governor Setting Command Specific Status Values ..... 26
Table 25: C8h - Set/ Get SMBUS Address Setting – Command Dword 11 ..... 26
Table 26: C9h – Set/Get LED pattern or pin11 pattern ..... 27
Table 27: D5h – Reset Timed Workload Counters – Command Dword 11 ..... 27
Table 28: E2h – Set/Get Enable Latency Tracking ..... 27
Table 29: NVMe Driver Support ..... 28
Table 30: Device Certifications and Declarations ..... 29
Table 31: Identify Controller ..... 30
Table 32: Power State Descriptor ..... 34
Table 33: Identify Namespace ..... 35
Table 34: LBA Format Data Structure ..... 37
Table 35: Vital Product Data Structure (VPD) ..... 38
Table 36: Capability List Pointer (Out of Band Temperature Sensor) ..... 38
Table 37: Register 0x05 read out format ..... 39
Table 38: Subsystem Management Data Structure ..... 40
Table 39: Command Response on 0x6A and 0x6B (Intel Specific) ..... 43
Table 40: PCIe ID ..... 44
Table 41: LED Functionality (Preliminary Proposal) ..... 46

# 1 Overview

This document describes the specifications and capabilities of the Intel® Solid State Drive (SSD) DC P3608 Series.

Based on the Intel 20nm HET (High Endurance Technology) NAND, the Intel® SSD DC P3608 Series is the highest capacity and performance in the family of SSDs. It is available in the low profile PCIe® 3.0 x8 add-in-card at capacities of 1.6TB, 3.2TB, and 4TB. The power requirement defaults for up to 40W[1], and can be set to 35W and as low as 25W for power sensitive systems. The real world power varies depending on workload and is the highest during primarily sequential write applications. The Intel® SSD DC P3608 Series features the same data center RAS features as the other drives in the family, such as:

**Enhanced PLI (Power Loss Imminent)** – protection from unplanned power loss obtained by a propriety combination of hardware, firmware algorithms, and the robust validation at a multitude of corner cases.

**Intel 20nm HET (High Endurance Technology) NAND** for up to 3 DWPD (Drive Writes Per Day) over 5 years. The 4TB can write up to 21.9PB over the life of the drive[2].

**End-to-end data protection** and ECC on all internal and external memories in the data path for protection at every layer.

**Dual NVMe Controller** – The NVMe protocol is more efficient than SATA or SAS providing more IOPS at a lower CPU utilization[3]. The dual controller architecture allows the SSD to simultaneously transfer data with this improved efficiency, and can be aggregated through the Intel® RSTe software.

**NVMe™ SMART monitoring and sideband health monitoring** – monitor the health of the SSD for attributes like temperature, endurance, power cycles, power on hours, and NVMe critical warnings. Sideband or out-of-band health monitoring enables OS agnostic monitoring through the standard NVM Express Management Interface Specification.

NVMe brings high performance and low latency PCIe SSDs into the mainstream with a streamlined protocol that is efficient and scalable, with the ease of deployment through industry standard software and drivers. Joining the family is the Intel® SSD DC P3608 Series - expanding the maximum capacity to 4TB, doubling the interface bandwidth to obtain real world transfers of over 5GB/s, with a unique NVMe dual-controller architecture to achieve maximum storage efficiency.

1, 2, 3: 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. For more complete information about performance and benchmark results, visit http://www.intel.com/performance. Tests performed by Intel. Test and System Configurations: Intel® Core™ i7-4770 CPU at 3.4 GHz, 8GB DDR3 at 1600 MHz, Intel® SSD DC P3608 Series 1.6TB.

## 1.1 References

**Table 1: Standard Information Referenced in this Document**

| Date | Title | Location |
|---|---|---|
| Jan 2013 | Enterprise SSD Form Factor Version 1.0a | http://www.ssdformfactor.org |
| Feb 2012 | NVM Express Revision 1.0 | http://www.nvmexpress.org |
| Nov 2010 | PCIe Base Specification Revision 3.0 | http://pcisig.com |
| July 2012 | Solid-State Drive (SSD) Requirements and Endurance Test Method (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/<br>Click *Search MDDS Database* and search for material description datasheet |
| 1995<br>1996<br>1995<br>1995<br>1997<br>1994 | International Electrotechnical Commission EN 61000<br>4-2 (Electrostatic discharge immunity test)<br>4-3 (Radiated, radio-frequency, electromagnetic field immunity test)<br>4-4 (Electrical fast transient/burst immunity test)<br>4-5 (Surge immunity test)<br>4-6 (Immunity to conducted disturbances, induced by radio-frequency fields)<br>4-11 (Voltage Variations, voltage dips, short interruptions and voltage variations immunity tests) | http://www.iec.ch/ |
| 1995 | ENV 50204<br>(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 |
| EXT | Extended |
| FPDMA | First Party Direct Memory Access |
| GB | Gigabyte<br>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 NAND flash management and maintenance purposes. |
| Gb | Gigabit |
| HDD | Hard Disk Drive |
| HET | High Endurance Technology |
| KB | Kilobyte |
| I/O | Input/Output |
| IOMeter | I/O Subsystem Measurement Tool |
| IOPS | Input/Output Operations Per Second |
| ISO | International Standards Organization |
| LBA | Logical Block Address |
| MB | Megabyte (1,000,000 bytes) |
| MLC | Multi-level Cell |
| 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 |
| SSD | Solid-State Drive |
| TB | Terabyte |
| TYP | Typical |
| UBER | Uncorrectable Bit Error Rate |
| VPD | Vital Product Data |

§

# 2 Product Specifications

## 2.1 Capacity

**Table 3: User Addressable Sectors**

| Intel SSD DC P3608 Series | Unformatted Capacity (Total User Addressable Sectors in LBA Mode) |
|---|---|
| 1.6TB | 3,125,627,568 |
| 3.2TB | 6,251,233,968 |
| 4.0TB | 7,814,037,168 |

***Note:*** 1 GB = 1,000,000,000 bytes; 1 sector = 512 bytes or 520 bytes or 528 bytes. LBA count shown represents total user storage capacity and will remain the same throughout the life of the drive. 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.

## 2.2 Performance

**Table 4: Random Read/Write Input/Output Operations Per Second (IOPS)**

| Specification[1] | Unit | Intel SSD DC P3608 Series[1] | | |
|---|---|---|---|---|
| | | 1.6TB[4] | 3.2TB[4] | 4.0TB[4] |
| Random 4KB 70/30 Read/Write (up to)[2] | IOPS | 300,000 | 165,000 | 150,000 |
| Random 8KB 70/30 Read/Write (up to)[3] | IOPS | 160,000 | 80,000 | 75,000 |
| Random 4KB Read (up to) | IOPS | 850,000 | 850,000 | 850,000 |
| Random 4KB Write (up to) | IOPS | 150,000 | 80,000 | 50,000 |
| Random 8KB Read (up to) | IOPS | 500,000 | 500,000 | 500,000 |
| Random 8KB Write (up to) | IOPS | 60,000 | 36,000 | 28,000 |

1. Performance measured using IOMeter* on Intel provided Windows Server 2012 R2 driver with Queue Depth 256 and number of workers equal to 8. Measurements are performed on a full Logical Block Address (LBA) span of the drive. Power mode set at 40 W. Values are aggregate over both partitions with equal workloads.
2. 4KB = 4,096 bytes
3. 8KB = 8,192 bytes
4. 1TB = $10^{12}$ bytes

### Table 5: Random Read/Write IOPS Consistency

| Specification[1] | Unit | Intel SSD DC P3608 Series | | |
|---|---|---|---|---|
| | | 1.6TB | 3.2TB | 4.0 TB |
| Random 4KB Read (up to)[2] | % | 90 | 90 | 90 |
| Random 4KB Write (up to) | % | 90 | 90 | 90 |
| Random 8KB Read (up to)[3] | % | 90 | 90 | 90 |
| Random 8KB Write (up to) | % | 90 | 90 | 90 |

1. Performance consistency measured using IOMeter based on Random 4KB with total queue depth of 128, measured as (IOPS in the 99.9th percentile slowest 1-second interval)/(average IOPS during the test). Measurements are performed on a full Logical Block Address (LBA) span of the drive once the workload has reached steady state but including all background activities required for normal operation and data reliability
2. 4KB = 4,096 bytes
3. 8KB = 8,192 bytes

### Table 6: Sequential Read and Write Bandwidth

| Specification | Unit | Intel SSD DC P3608 Series | | |
|---|---|---|---|---|
| | | 1.6TB | 3.2TB | 4.0TB |
| Sequential Read (up to)[1] | MB/s | 5,000 | 4,500 | 5,000 |
| Sequential Write (up to)[1] | MB/s | 2,000 | 2,600 | 3,000 |

**NOTE:**

1. Performance measured using IOMeter with 128 KB (131,072 bytes) of transfer size with Queue Depth 128. Power mode set at 40W.

### Table 7: Latency

| Specification | Intel SSD DC P3608 Series |
|---|---|
| | 1.6TB, 3.2TB, 4.0TB |
| Latency[1] (TYP)<br>Read Sequential/Random<br>Write Sequential/Random<br>Power On to PCIe Config Ready[2]<br>Power On to Controller Ready[2] | <br>20/120 µs<br>20/30 µs<br>2.0 sec (TYP)<br>10.0 sec (TYP) |

1. Latency measured using 4 KB (4,096 bytes) transfer size with Queue Depth equal to 1 using Windows Server 2012 R2 driver. Power mode set at 40W.
2. Power On To Ready time measured from de-assertion of PCIe reset to PCIe Config Ready state.

   Power On to Controller ready 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 20 seconds.

**Table 8: Quality of Service**

| Specification | Unit | Intel SSD DC P3608 Series | |
|---|---|---|---|
| | | QD=1 | QD=128 |
| **Quality of Service[1,2] (99%)** | | | |
| Reads | ms | <0.120 | <0.750 |
| Writes | ms | <0.500 | <25 |
| **Quality of Service[1,2] (99.99%)** | | | |
| Reads | ms | <3 | <7 |
| Writes | ms | <4.7 | <30 |

1. Device measured using IOMeter. Quality of Service measured using 4KB (4,096 bytes) transfer size on a random workload on a full Logical Block Address (LBA) span of the drive once the workload has reached steady state but including all background activities required for normal operation and data reliability.
2. Based on Random 4KB QD=1,128 workloads, measured as the time taken for 99.0 (or 99.99) percentile of commands to finish the round-trip from host to drive and back to host.

## 2.3 Electrical Characteristics

**Table 9: Operating Voltage**

| Electrical Characteristics | Intel SSD DC P3608 Series |
|---|---|
| **3.3V Operating Characteristics:**[1] | |
| Operating Voltage range | 3.3 V (±10%) |
| Rise Slew Rate / Fall Slew Rate | 66V/s – 3.3kV/s / 660mV/s – 3.3kV/s |
| Startup Overshoot | 4.6V (26ms) |
| Glitch High | 4.6V, 40us |
| Noise level | 300 mV pp 10Hz – 100 KHz<br>50 mV pp 100KHz – 20 MHz |
| Min Off time | 3 seconds |
| Inrush Current (Typical Peak) | 1.5 A ($T_{startup}$ ≤ 2 Sec) |
| Max Average Current | 3.0 A |

| Electrical Characteristics | Intel SSD DC P3608 Series |
|---|---|
| **12V Operating Characteristics:**[1] | |
| Operating Voltage range | 12 V (+10%/-20%) |
| Rise Slew Rate / Fall Slew Rate | 240V/s – 12kV/s / 2.4V/s - 12kV/s |
| Startup Overshoot | 15.0V for 26ms |
| Glitch High | 15.0V, 40us |
| Noise level | 1000 mV pp 10Hz – 100 KHz<br>100 mV pp 1MHz – 20 MHz |
| Min Off time | 3 seconds |
| Inrush Current (Typical Peak) | 1.5A ($T_{startup} \leq$ 2sec) |
| Max Average Current | 2.1A for 25W slot / 5.5A for 75W slot |
| **3.3Vaux Operating Characteristics:**[2] | |
| Operating Voltage range | 3.3V (±9%) |
| Rise Slew Rate / Fall Slew Rate | 250V/s – 420kV/s >0.66V/s (5 second) |
| Startup Overshoot | 4.6V (26ms) |
| Glitch Low / Glitch High | 2.5V for 30us / 4.6V for 40us |
| Noise level | 300 mV pp 10Hz – 100 KHz<br>50 mV pp 100KHz – 20 MHz |
| Max Current | 50mA |

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. 3.3Vaux is optional, not needed for power up or functionality. 3.3Vaux is needed for accessing VPD page by means of SMBUS for both form factors.

### Table 10: Power Consumption

| Specification | Unit | Intel SSD DC P3608 Series | | |
|---|---|---|---|---|
| | | 1.6TB | 3.2TB | 4.0TB |
| Active Write - Average [1] | W | 30 | 35 | 40 |
| Active Read - Average [2] | W | 18 | 18 | 20 |
| Idle | W | 11.5 | 11.5 | 11.5 |

1. The workload equates 128KB (131,072 bytes) Queue Depth equal to 128 sequential writes. Average power is measured using scope trigger over a 100 ms sample period
2. The workload equates 128KB (131,072 bytes) Queue Depth equal to 128 sequential reads.

## 2.4 Environmental Conditions

**Table 11: Temperature, Shock, Vibration**

| Temperature | Add-In Card form factor |
|---|---|
| Temperature<br>Operating[1]<br>Non-operating[3] | <br>Ambient 0 – 55° C / 0 -40° C[2]<br>-55–95° C |
| Temperature Gradient[4]<br>Operating<br>Non-operating | <br>30° C/hr (Typical)<br>30° C/hr (Typical) |
| Humidity<br>Operating<br>Non-operating | <br>5–95%<br>5–95% |
| **Shock and Vibration** | **Range** |
| Shock[5]<br>Operating<br>Non-operating | <br>50 G Trapezoidal, 170 in/s<br>50 G Trapezoidal, 170 in/s |
| Vibration[6]<br>Operating<br>Non-operating | <br>2.17 GRMS (5-700 Hz) Max<br>3.13 GRMS (5-800 Hz) Max |

1. Operating temperature implies ambient air temperature under defined airflow in Tables 12 and 13
2. 0-55° C is for airflow from the server towards the card and 0-40° C is for airflow into the server
3. Please contact your Intel representative for details on the non-operating temperature range
4. Temperature gradient measured without condensation
5. 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
6. 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

**Table 12: Airflow Requirements for Intel SSD DC P3608 Series (Add-In Card)**

| Airflow Direction | Unit | Ambient Temperature | Intel SSD DC P3608 Series | | |
|---|---|---|---|---|---|
| | | | 1.6TB | 3.2TB | 4.0TB |
| Into Server | | 0 - 40° C | 400 | 400 | 400 |

## 2.5 Product Regulatory Compliance

Intel SSD DC P3608 Series meets or exceeds the regulatory or certification requirements in the following table.

**Table 13: Product Regulatory Compliance Specifications**

| Title | Description | Region For Which Conformity Declared |
|---|---|---|
| TITLE 47-Telecommunications CHAPTER 1 – FEDERAL COMMUNMICATIONS COMMISSION PART 15 – RADIO FREQUENCY DEVICES | FCC Part 15B Class A | USA |
| ICES-003, Issue 4 Interference-Causing Equipment Standard Digital Apparatus | CA/CSA-CEI/IEC CISPR 22:10. This is CISPR 22:2008 with Canadian Modifications | Canada |
| IEC 55024 Information Technology Equipment – Immunity characteristics – Limits and methods of measurement CISPR24:2010 | EN-55024: 2010 and its amendments | European Union |
| IEC 55022 Information Technology Equipment – Radio disturbance Characteristics – Limits and methods of measurement CISPR24:2008 (Modified) | EN-55022: 2010 and its amendments | European Union |
| EN-60950-1 2nd Edition | Information Technology Equipment – Safety – Part 1: General Requirements | USA/Canada |
| UL/CSA EN-60950-1 2nd Edition | Information Technology Equipment – Safety – Part 1: General Requirements | USA/Canada |

## 2.6 Reliability Specifications

Intel SSD DC P3608 Series meets or exceeds SSD endurance and data retention requirements as specified in the JESD218 standard. Reliability specifications are listed in the table below.

**Table 14: Reliability Specifications**

| Parameter | Value |
|---|---|
| Uncorrectable Bit Error Rate (UBER)<br>Uncorrectable bit error rate will not exceed one sector in the specified number of bits read. In the unlikely event of a 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)<br>Mean Time Between Failures is estimated based on Telcordia* methodology and demonstrated through Reliability Demonstration Test (RDT). | >= 1 million hours |
| Data Retention<br>The time period for retaining data in the NAND at maximum rated endurance. | 3 months power-off retention once SSD reaches rated write endurance at 40° C |
| Endurance Rating<br>The number of drive writes such that the SSD meets the requirements according to the JESD218 standard. Endurance rating verification is defined to establish UBER <1E-16 at 60% upper confidence limit. | 1.6TB: 8.76 PBW<br>3.2TB: 17.52 PBW<br>4.0TB: 21.90 PBW<br>(3 drive writes/day*) |

**NOTE:** Petabytes Written (PBW). Refer to JESD218 standard table 1 for UBER, FFR and other Enterprise SSD requirements.

## 2.7 Temperature Sensor

Intel® SSD DC P3608 Series has a total of 4 temperature sensors with 1 near each controller ASIC and one internal to each EEPROM, all with an accuracy of +/-2° C over a range of -10° C to +85° C. These can be monitored using the NVMe Health Log functionality. For more information on sensor reading, see the SMART attributes section.

In addition, the drive provides out of band reading of all temperatures on SMBUS as follows:

- Slave Address 6Ah to read Composite Temperature from SMART data pertaining to ASIC1, as defined in Appendix D. This requires 12V be applied to SSD
- Slave Address 6Bh to read Composite Temperature from SMART data pertaining to ASIC2, as defined in Appendix D. This requires 12V be applied to SSD
- Slave address 1Bh to read temperature from VPD EEPROM close to ASIC1. This requires 3.3Vaux supply be applied to SSD
- Slave address 19h to read temperature from VPD EEPROM close to ASIC2. This requires 3.3Vaux supply be applied to SSD

## 2.8 Power Loss Capacitor Test

Intel® SSD DC P3608 Series support testing of the power loss capacitor. Each discrete module has its own PLI cap and its own individual SMART logs (module #1 and module #2). The drive supports full data protection during power-loss interrupts, via discrete electrolytic capacitors. The health of these capacitors are self-tested at every power on and at regular frequency throughout the life of the drive. SMART Critical Warning Bit 4/Byte 0 are set if these capacitors are detected bad during the capacitor self-test, indicating the replacement of the drive is necessary.

## 2.10 Out of Band Management (SMBUS)

Intel® SSD DC P3608 Series provides out of band management by means of SMBUS interface. This requires 3.3V Auxiliary voltage. SMBUS accesses a VPD page as listed in Appendix B through address 0x53/0xA6 (7bit/8bit).

Temperature sensor is accessed through address 0x1B/0x36(7bit/8bit). For temperature sensor access, temperature can be read by the BMC (base) using Read Temperature Data Register (0x05) by means of SMBUS 0x1B. Bits [11:0] return raw ambient temperature.

Intel SSD DC P3608 provides additional drive information via address 0x6A/0xD4(7bit/8bit) as outlined in the NVMe Basic Management Command (see http://www.nvmexpress.org/wp-content/uploads/NVMe_Management_-_Technical_Note_on_Basic_Management_Command.pdf).

Appendix B, C and D have details on the Out of Band Management data structure.

**NOTE:** In certain tools the address for the VPD and temperature sensor will appear as 0xA6 and 0x36 respectively, due to bit shift.

## 2.11 Variable Sector Size Support

P3608 Series supports vector sizes of 512, 520, 528, 4096, 4104, 4160 and 4224 bytes. P3608 Series will also support DIF as specified in NVMe 1.0 specification. 520 and 4104 Byte sector sizes can support PI (protection information) which is 8 Byte long.

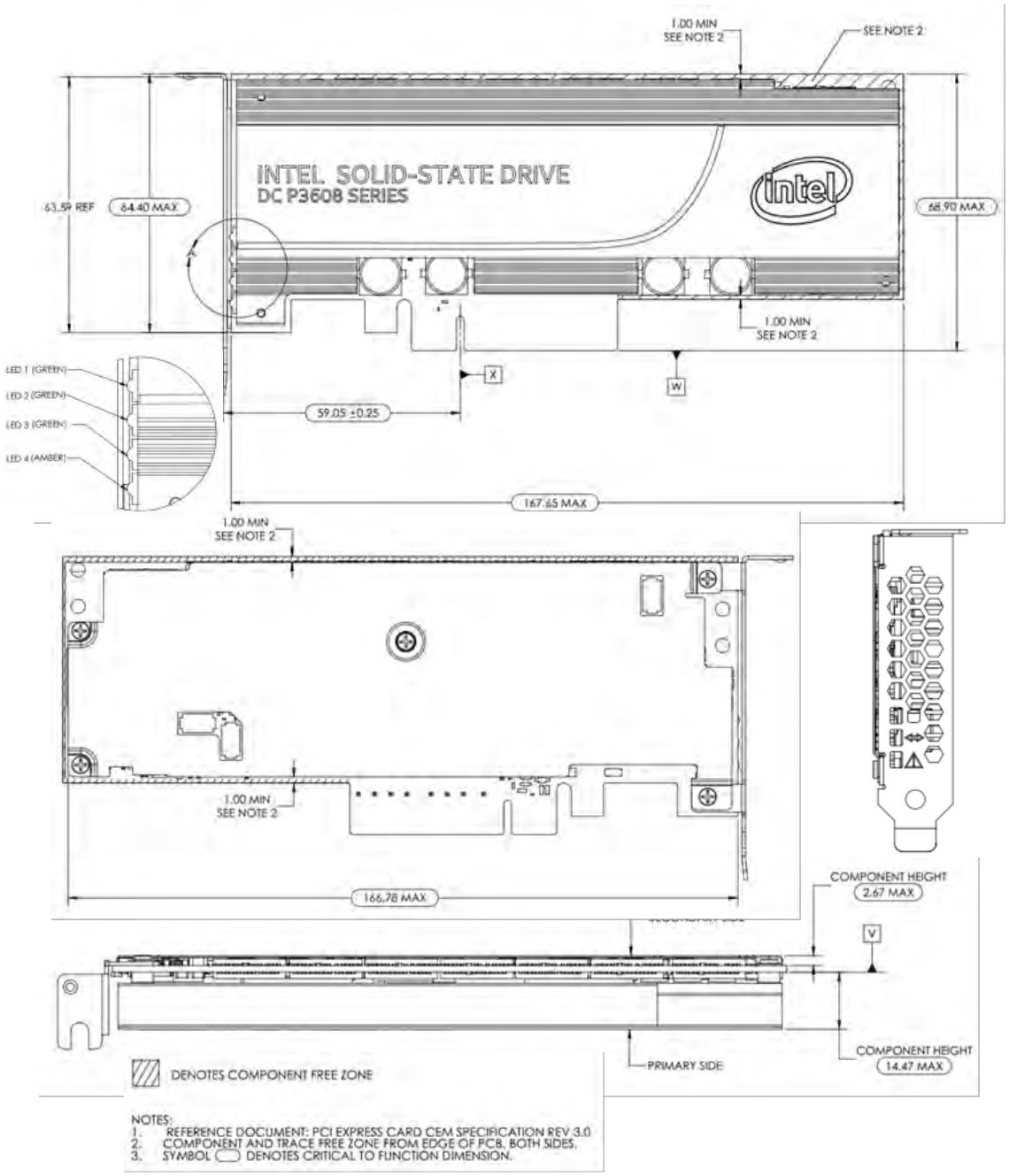
Device only supports PRACT=0, implying protection information is passed to the SSD and checked by the SSD.

§

# 3 Mechanical Information

Figure 3-1 shows the physical package information for the Intel SSD DC P3608 Series in the x8 AIC. All dimensions are in millimeters.

**Figure 3-1: Intel SSD DC P3608 Series Dimensions**




§

# 4 Pin and Signal Descriptions

## 4.1 AIC Pin Definition Table

**Table 15: Pin Definition for Add-In Card (Half Height Half Length) Form Factor**

| | Side B | | Side A | |
|---|---|---|---|---|
| **Pin** | **Name** | **Description** | **Name** | **Description** |
| 1 | +12V | 12V power | PRSNT1# | Hot-Plug presence detect |
| 2 | +12V | 12V power | +12V | 12V power |
| 3 | +12V | 12V power | +12V | 12V power |
| 4 | GND | Ground | GND | Ground |
| 5 | SMCLK | SMBus (System Management Bus) clock | JTAG2 | TCK (Test Clock), clock input for JTAG interface |
| 6 | SMDAT | SMBus (System Management Bus) data | JTAG3 | TDI (Test Data Input) |
| 7 | GND | Ground | JTAG4 | TDO (Test Data Output) |
| 8 | +3.3V | 3.3V power | JTAG5 | TMS (Test Mode Select) |
| 9 | JTAG1 | TRST# (Test Reset) resets the JTAG interface | +3.3V | 3.3V power |
| 10 | 3.3Vaux | 3.3V auxiliary power | +3.3V | 3.3V power |
| 11 | WAKE# | Signal for Link reactivation | PERST# | Fundamental reset |
| Mechanical Key | | | | |
| 12 | RSVD | Reserved | GND | Ground |
| 13 | GND | Ground | REFCLK+ | Reference clock (differential pair) |
| 14 | PETp0 | Transmitter differential pair, Lane 0 | REFCLK- | Reference clock (differential pair) |
| 15 | PETn0 | Transmitter differential pair, Lane 0 | GND | Ground |
| 16 | GND | Ground | PERp0 | Receiver differential pair, Lane 0 |
| 17 | PRSNT2# | Hot-Plug presence detect | PERn0 | Receiver differential pair, Lane 0 |
| 18 | GND | Ground | GND | Ground |
| 19 | PETp1 | Transmitter differential pair, Lane 1 | RSVD | Reserved |
| 20 | PETn1 | Transmitter differential pair, Lane 1 | GND | Ground |
| 21 | GND | Ground | PERp1 | Receiver differential pair, Lane 1 |
| 22 | GND | Ground | PERn1 | Receiver differential pair, Lane 1 |
| 23 | PETp2 | Transmitter differential pair, Lane 2 | GND | Ground |
| 24 | PETn2 | Transmitter differential pair, Lane 2 | GND | Ground |
| 25 | GND | Ground | PERp2 | Receiver differential pair, Lane 2 |

| | Side B | | Side A | |
|---|---|---|---|---|
| **Pin** | **Name** | **Description** | **Name** | **Description** |
| 26 | GND | Ground | PERn2 | Receiver differential pair, Lane 2 |
| 27 | PETp3 | Transmitter differential pair, Lane 3 | GND | Ground |
| 28 | PETn3 | Transmitter differential pair, Lane 3 | GND | Ground |
| 29 | GND | Ground | PERp3 | Receiver differential pair, Lane 3 |
| 30 | RSVD | Reserved | PERn3 | Receiver differential pair, Lane 3 |
| 31 | PRSNT2# | Hot-Plug presence detect | GND | Ground |
| 32 | GND | Ground | RSVD | Reserved |
| 33 | PETp4 | Transmitter differential pair, Lane 4 | RSVD | Reserved |
| 34 | PETn4 | Transmitter differential pair, Lane 4 | GND | Ground |
| 35 | GND | Ground | PERp4 | Receiver differential pair, Lane 4 |
| 36 | GND | Ground | PERn4 | Receiver differential pair, Lane 4 |
| 37 | PETp5 | Transmitter differential pair, Lane 5 | GND | Ground |
| 38 | PETn5 | Transmitter differential pair, Lane 5 | GND | Ground |
| 39 | GND | Ground | PERp5 | Receiver differential pair, Lane 5 |
| 40 | GND | Ground | PERn5 | Receiver differential pair, Lane 5 |
| 41 | PETp6 | Transmitter differential pair, Lane 6 | GND | Ground |
| 42 | PETn6 | Transmitter differential pair, Lane 6 | GND | Ground |
| 43 | GND | Ground | PERp6 | Receiver differential pair, Lane 6 |
| 44 | GND | Ground | PERn6 | Receiver differential pair, Lane 6 |
| 45 | PETp7 | Transmitter differential pair, Lane 7 | GND | Ground |
| 46 | PETn7 | Transmitter differential pair, Lane 7 | GND | Ground |
| 47 | GND | Ground | PERp7 | Receiver differential pair, Lane 7 |
| 48 | PRSNT2# | Hot-Plug presence detect | PERn7 | Receiver differential pair, Lane 7 |
| 49 | GND | UART2 | GND | Ground |

- All pins are numbered in ascending order from the left to the right, with side A on the top of the centerline and side B on the bottom of the centerline, use the reference drawing in Fig2, with the logo visible.
- The PCI Express interface pins PETpx, PETnx, PERpx, and PERnx are named with the following convention: “PE” stands for PCI Express high speed, “T” for Transmitter, “R” for Receiver, “p” for positive (+) and “n” for negative (-).

§

# 5 Supported Command Sets

Intel SSD DC P3608 Series supports all mandatory Admin and I/O commands defined in NVMe™ (Non-Volatile Memory Express™) revision 1.0.

## 5.1 NVMe Admin Command Set

P3608 Series 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 Notification

P3608 Series also supports the following optional I/O commands defined in NVMe revision 1.0:

- Firmware Activate
- Firmware Image Download
- Format NVM

***Note:*** See Appendix A, “Identify Controller Data Structure” for details on commands and capabilities.

## 5.2 NVMe I/O Command Set

P3608 Series supports all the mandatory NVMe™ I/O command set defined in NVMe 1.0 specification, which are:

- Flush
- Write
- Read

Additionally, the following optional commands are supported:

- Write Uncorrectable
- Dataset Management (De-allocate only)

## 5.3 Log Page Support

Intel SSD DC P3608 Series supports the following mandatory log pages defined in NVMe 1.0 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 1.0 version of the specification for the log page content. Additionally, P3608 Series will support the following vendor unique log pages:

- Log Page Directory (Log Identifier C0h)
- Read Command Latency Statistics Log (Log Identifier C1h)
- Write Command Latency Statistics Log (Log Identifier C2h)
- Temperature Statistics (Log Identifier C5h)
- Vendor Unique SMART Log (Log Identifier CAh)
- Drive Marketing Name Log (Log Identifier DDh)

5.4 SMART Attributes

The following table lists the SMART attributes supported by the P3608 Series in accordance with NVMe 1.0 specification.

**Table 16: SMART Attributes (Log Identifier 02h)**

| Byte | # of Bytes | Attribute | Description |
|---|---|---|---|
| 0 | 1 | **Critical Warning**: These bits if set, flag various warning sources.<br>Bit 0: Available Spare is below Threshold<br>Bit 1: Temperature has exceeded Threshold<br>Bit 2: Reliability is degraded due to excessive media or internal errors<br>Bit 3: Media is placed in Read- Only Mode<br>Bit 4: Volatile Memory Backup System has failed (e.g., enhanced power loss capacitor test failure)<br>Bits 5-7: Reserved | Any of the critical warning can be tied to asynchronous event notification. Drive Health Indicator defined under bytes 3095-3076 of Identify Controller may still indicate “healthy” status when the critical warning flag is set. |
| 1 | 2 | Temperature: Overall Device current temperature in Kelvin. | For AIC, it reports the NAND temperature. |
| 3 | 1 | Available Spare: Contains a normalized percentage (0 to 100%) of the remaining spare capacity available | Starts from 100 and decrements. |
| 4 | 1 | Available Spare Threshold | Threshold is set to 10%. |
| 5 | 1 | Percentage Used Estimate<br>(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).<br><br>If the value reaches or exceeds 105 (5% above rated endurance level), the drive may be performing writes at very low rate. |

| Byte | # of Bytes | Attribute | Description |
|---|---|---|---|
| 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 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. |

**Table 17: Additional SMART Attributes (Log Identifier CAh)**

| Byte | # of Bytes | Attribute | Description |
|---|---|---|---|
| 0 | 1 | AB (Program Fail Count) | Raw value: shows total count of program fails.<br>Normalized value: beginning at 100, shows the percent remaining of allowable program fails. |
| 3 | 1 | Normalized Value | |
| 5 | 6 | Current Raw Value | |
| 12 | 1 | AC (Erase Fail Count) | Raw value: shows total count of erase fails.<br>Normalized value: beginning at 100, shows the percent remaining of allowable erase fails. |
| 15 | 1 | Normalized Value | |
| 17 | 6 | Current Raw Value | |
| 24 | 1 | AD (Wear Leveling Count) | Raw value:<br>Bytes 1-0: Min. erase cycle<br>Bytes 3-2: Max. erase cycle<br>Bytes 5-4: Avg. erase cycles<br>Normalized value: decrements from 100 to 0. |
| 27 | 1 | Normalized Value | |
| 29 | 6 | Current Raw Value | |
| 36 | 1 | B8 (End to End Error Detection Count) | Raw value: reports number of End-to-End detected and corrected errors by hardware.<br>Normalized value: always 100. |
| 39 | 1 | Normalized Value | |
| 41 | 6 | Current Raw Value | |
| 48 | 1 | C7 (CRC Error Count) | Raw value: shows total number of PCIe Interface CRC errors encountered, as specified in PCIe Link Performance Counter Parameter for “Bad TLP”.<br>Normalized value: always 100. |
| 51 | 1 | Normalized Value | |
| 53 | 6 | Current Raw Value | |
| 60 | 1 | 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.<br>Normalized value: always 100. |
| 63 | 1 | Normalized Value | |
| 65 | 6 | Current Raw Value | |
| 72 | 1 | 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.<br>Normalized value: always 100. |
| 75 | 1 | Normalized Value | |
| 77 | 6 | Current Raw Value | |
| 84 | 1 | E4 (Timed Workload, Timer) | Raw value: measures the elapsed time (number of minutes since starting this workload timer).<br>Normalized value: always 100. |
| 87 | 1 | Normalized Value | |
| 89 | 6 | Current Raw Value | |

| Byte | # of Bytes | Attribute | Description |
|---|---|---|---|
| 96 | 1 | EA (Thermal Throttle Status) | Raw value: reports Percent Throttle Status and Count of events<br>Byte 0: Throttle status reported as integer percentage.<br>Bytes 1-4: Throttling event count. Number of times thermal throttle has activated. Preserved over power cycles.<br>Byte 5: Reserved.<br>Normalized value: always 100. |
| 99 | 1 | Normalized Value | |
| 101 | 6 | Current Raw Value | |
| 108 | 1 | F0 (Retry Buffer Overflow Counter) | Raw Value: Counter to indicate the number of times Retry Buffer has overflown<br>Normalized Value is always 100 |
| 111 | 1 | Normalized Value | |
| 113 | 6 | Current Raw Value | |
| 120 | 1 | F3 (PLL Lock Loss Count) | Raw Value: Counter to indicate the number of times PCIe Refclock PLL has unlocked<br>Normalized Value is always 100 |
| 123 | 1 | Normalized Value | |
| 125 | 6 | Current Raw Value | |
| 132 | 1 | F4 (NAND Bytes Written) | Nand sectors written divided by 65536 (1 count = 32 MiB)<br>Normalized value always 100 |
| 135 | 1 | Normalized Value | |
| 137 | 6 | Current Value | |
| 144 | 1 | F5 (Host Bytes Written) | Host sectors written divided by 65536 (1 count = 32 MiB)<br>Normalized value always 100 |
| 147 | 6 | Normalized Value | |
| 149 | 6 | Current Value | |

## 5.5 Temperature Statistics

**Table 18: Temperature Statistics (Log Identifier C5h)**

| Byte | # of Bytes | Description |
|---|---|---|
| 0 | 1 | Current Temperature |
| 8 | 8 | Overtemp shutdown Flag for last critical component temperature |
| 16 | 8 | Overtemp shutdown Flag for life critical component temperature |
| 24 | 8 | Highest temperature |
| 32 | 8 | Lowest temperature |
| 80 | 8 | Specified Maximum Operating Temperature |
| 96 | 8 | Specified Minimum Operating Temperature |
| 104 | 8 | Estimated Offset |

## 5.6 Drive Marketing Name Log

**Table 19: 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 |
| 16 | 5 | P3608 |
| 21 | 1 | Space |
| 22 | 6 | Series |
| 28-511 | 484 | Reserved |

## 5.7 IO Command Latency Statistics

**Table 20: Read/Write Command Latency Log (Log Identifier C1h/C2h)**

| Byte | # of Bytes | Log Page Content |
|---|---|---|
| 0 | 2 | Major Revision |
| 2 | 2 | Minor Revision |
| 4 | 128 | First group of buckets: range 0-1ms, step 32us, each bucket size is 4 bytes, total 32 buckets |
| 132 | 124 | Second group of buckets: range 1-32ms, step 1ms, each bucket size is 4 bytes, total 31 buckets |
| 256 | 124 | Third group of buckets: range 32ms-1s, step 32ms, each bucket size is 4 bytes, total 31 buckets |

## 5.7 SET Feature Identifiers

In addition to the SMART attribute structure, features pertaining to the operation and health of the Intel SSD DC P3608 Series can be reported to the host on request through the Get Features command. P3608 Series can change settings using SET Features on the following items as defined in NVMe 1.0 specification.

- 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)

Intel SSD DC P3608 Series will also support the following vendor unique SET Features.

- Set/Get Max LBA (Feature Identifier C1h)
- Set/Get Native Max LBA (Feature Identifier C2h)
- Power Governor Setting (Feature Identifier C6h)
- Reset Timed Workload Counters (Feature Identifier D5h)

**Table 21: 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.<br>Read Usage: This field contains the 64-bit maximum LBA addressable by the Drive.<br>Command Dword 11 contains bits 31:00; Command Dword 12 contains bits 63: 32. |

**Table 22: 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 23: C6h - Set/ Get Power (Typical) Governor Setting – Command Dword 11**

| Bit | Description |
|---|---|
| 31:08 | Reserved (TBD) |
| 07:00 | Power Governor Setting: 00h = 40W , 01h = 35W , 02h = 25W |

**Table 24: Status Codes - Power Governor Setting Command Specific Status Values**

| Value | Description |
|---|---|
| C0h | Invalid Setting |

**Table 25: C8h - Set/ Get SMBUS Address Setting – Command Dword 11**

| Bit | Description |
|---|---|
| 0 | Reserved (TBD) |
| 08:01 | SMBUS controller address |
| 31:09 | Reserved |

***Note:***
Any address equal to or greater than 7Fh will disable the SMBUS address

### Table 26: C9h – Set/Get LED pattern or pin11 pattern

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

***Note:***

C9- Set Features Command Dword 11 is divided into the 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 is 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]

### Table 27: 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 Table 25.

### Table 28: E2h – Set/Get Enable Latency Tracking

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

§

# 6 NVMe Driver Support

The following table describes the NVMe Driver Support for Intel SSD DC P3608 Series. 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 29: NVMe Driver Support**

| Support Level | Operating System Description |
|---|---|
| Intel Provided[1] | Windows* Server 2012 R2, 2012, 2008 R2 x64, Windows 7(32bit/64bit), Windows 8 (32bit/64bit), Windows 8.1 (32bit/64Bit) |
| In-box Driver[2] | RHEL 6.5, RHEL 6.6, RHEL6.7, RHEL 7.0, RHEL 7.1, SLES11 SP4, Windows* Server 2012 R2, Windows 8.1, Citrix Xenserver 6.5, Cent OS 6.5 |

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

§

# 7 Certifications and Declarations

**Table 30: Device Certifications and Declarations**

| Certification | Description |
|---|---|
| CE Compliant | European Economic Area (EEA): Compliance with the essential requirements of EC Council Directives Low Voltage Directive (LVD) 2006/95/EC, EMC Directive 2004/108/EC and Directive 2011/65/EU. |
| UL Recognized | Certified Underwriters Laboratories, Inc. Bi-National Component Recognition; UL 60950-1, 2nd Edition, 2007-03-27 (Information Technology Equipment - Safety - Part 1: General Requirements)<br>CSA C22.2 No. 60950-1-07, 2nd Edition, 2007-03 (Information Technology Equipment - Safety - Part 1: General Requirements) |
| C-Tick Compliant | Compliance with the Australia/New Zealand Standard AS/NZS3548 and Electromagnetic Compatibility (EMC) Framework requirements of the Australian Communication Authority (ACA). |
| BSMI Compliant | 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. |
| 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. |
| VCCI | Voluntary Control Council for Interface to cope with disturbance problems caused by personal computers or facsimile. |
| Microsoft WHCK | Microsoft Windows Hardware Certification Kit |
| RoHS Compliant | Restriction of Hazardous Substance Directive |
| WEEE | Directive on Waste Electrical and Electronic Equipment |

§

# Appendix A IDENTIFY Data Structure

**Table 31: Identify Controller**

| Bytes | F = Fixed<br>V = Variable<br>X = Both | Default Value | Interpretation | Description |
|---|---|---|---|---|
| 1-0 | F | 8086h | Contains the company vendor identifier that is assigned by the PCI SIG | PCI Vendor ID (VID) |
| 3-2 | F | 8086h | Contains the company vendor identifier that is assigned by the PCI SIG for subsystem | PCI Subsystem Vendor ID (SSVID) |
| 23-4 | V | varies | Contains the serial number for the NVM subsystem | Serial Number (SN) |
| 63-24 | V | varies | Contains the serial number for the NVM subsystem that is assigned by the vendor as an ASCII string | Model Number (MN) |
| 71-64 | V | varies | Contains the currently active firmware revision for the NVM subsystem | Firmware Revision (FR) |
| 72 | F | 0h | Recommended Arbitration Burst size equals 1 | Recommended Arbitration Burst (RAB) |
| 75-73 | F | 5CD2E4h | Contains the Organization Unique Identifier (OUI) for the controller vendor | IEEE OUI Identifier (IEEE) |
| 76 | X | 0h | No of multiple PCI Express interfaces connected to the host, bit 0 determines multiple interface | Multi-Interface Capabilities (MIC) |
| 77 | F | 05h | Supports MDTS of 128K | Maximum Data Transfer Size (MDTS) |
| 255:78 | | | | Reserved |
| 257-256 | F | 06h | Supports Security Send/Receive, Format NVM and Firmware Activate/Download | Optional Admin Command Support (OACS) |
| 258 | F | 03h | Supports up to 3 concurrently outstanding abort commands | Abort Command Limit (ACL) |
| 259 | F | 03h | Supports up to 3 concurrently outstanding asynchronous event requests | Asynchronous Event Request Limit (AERL) |
| 260 | X | 02h | Single slot<br>Read/write capable | Firmware Updates (FRMW) |
| 261 | X | 02h | SMART/Health Log Support per drive not per namespace | Log Page Attributes (LPA) |
| 262 | F | 3Fh | Number of Error Information log entries equals 64 | Error Log Page Entries (ELPE) |
| 263 | F | 0h | Number of NVM Express power states equal 1 | Number of Power States Support (NPSS) |
| 264 | F | 0h | Configuration settings for Admin Vendor Specific command handling | Admin Vendor Specific Command Configuration (AVSCC) |
| 511-265 | | | | Reserved |

| Bytes | F = Fixed<br>V = Variable<br>X = Both | Default Value | Interpretation | Description |
|---|---|---|---|---|
| 512 | F | 66h | Required and max submission queue entry size is 64 Byte | Submission Queue Entry Size (SQES) |
| 513 | F | 44h | Required and max submission queue entry size is 16 Byte | Completion Queue Entry Size (CQES) |
| 515-514 | | | | Reserved |
| 519-516 | F | 01h | Supports single namespace | Number of Namespaces (NN) |
| 521-520 | F | 06h | Supports Dataset Management and Write Uncorrectable optional NVMe commands. | Optional NVMe Command Support (ONCS) |
| 523-522 | F | 0h | Fused commands not supported | Fused Operation Support (FUSES) |
| 524 | F | 07h | Supports Crypto Erase and format of entire drive, not per namespace | Format NVM Attributes (FNA): |
| 525 | F | 0h | Volatile write cache is not present | Volatile Write Cache (VWC) |
| 527-526 | F | 0h | Atomic write size for controller during normal equals to 512B | Atomic Write Unit Normal (AWUN) |
| 529-528 | F | 0h | Indicates the atomic write size for the controller during a power fail condition equals 512B | Atomic Write Unit Power Fail (AWUPF) |
| 530 | X | 0h | Not Supported | NVM Vendor Specific Command Configuration (NVSCC) |
| 703-531 | | | | Reserved |
| 2047-704 | | | | Reserved |
| 2079-2048 | V | | Indicates the characteristics of power state 0 | Power State 0 Descriptor (PSD0) |
| 2111-2080 | V | | Indicates the characteristics of power state 1 | Power State 1 Descriptor (PSD1) |
| 2143-2112 | V | | Indicates the characteristics of power state 2 | Power State 2 Descriptor (PSD2) |
| 2175-2144 | V | | Indicates the characteristics of power state 3 | Power State 3 Descriptor (PSD3) |
| 2207-2176 | V | | Indicates the characteristics of power state 4 | Power State 4 Descriptor (PSD4) |
| 2239-2208 | V | | Indicates the characteristics of power state 5 | Power State 5 Descriptor (PSD5) |
| 2271-2240 | V | | Indicates the characteristics of power state 6 | Power State 6 Descriptor (PSD6) |
| 2303-2272 | V | | Indicates the characteristics of power state 7 | Power State 7 Descriptor (PSD7) |
| 2335-2304 | V | | Indicates the characteristics of power state 8 | Power State 8 Descriptor (PSD8) |
| 2367-2336 | V | | Indicates the characteristics of power state 9 | Power State 9 Descriptor (PSD9) |
| 2399-2368 | V | | Indicates the characteristics of power state 10 | Power State 10 Descriptor (PSD10) |

| Bytes | F = Fixed<br>V = Variable<br>X = Both | Default Value | Interpretation | Description |
|---|---|---|---|---|
| 2431-2400 | V | | Indicates the characteristics of power state 11 | Power State 11 Descriptor (PSD11) |
| 2463-2432 | V | | Indicates the characteristics of power state 12 | Power State 12 Descriptor (PSD12) |
| 2495-2464 | V | | Indicates the characteristics of power state 13 | Power State 13 Descriptor (PSD13) |
| 2527-2496 | V | | Indicates the characteristics of power state 14 | Power State 14 Descriptor (PSD14) |
| 2559-2528 | V | | Indicates the characteristics of power state 15 | Power State 15 Descriptor (PSD15) |
| 2591-2560 | V | | Indicates the characteristics of power state 16 | Power State 16 Descriptor (PSD16) |
| 2623-2592 | V | | Indicates the characteristics of power state 17 | Power State 17 Descriptor (PSD17) |
| 2655-2624 | V | | Indicates the characteristics of power state 18 | Power State 18 Descriptor (PSD18) |
| 2687-2656 | V | | Indicates the characteristics of power state 19 | Power State 19 Descriptor (PSD19) |
| 2719-2688 | V | | Indicates the characteristics of power state 20 | Power State 20 Descriptor (PSD20) |
| 2751-2720 | V | | Indicates the characteristics of power state 21 | Power State 21 Descriptor (PSD21) |
| 2783-2752 | V | | Indicates the characteristics of power state 22 | Power State 22 Descriptor (PSD22) |
| 2815-2784 | V | | Indicates the characteristics of power state 23 | Power State 23 Descriptor (PSD23) |
| 2847-2816 | V | | Indicates the characteristics of power state 24 | Power State 24 Descriptor (PSD24) |
| 2879-2848 | V | | Indicates the characteristics of power state 25 | Power State 25 Descriptor (PSD25) |
| 2911-2880 | V | | Indicates the characteristics of power state 26 | Power State 26 Descriptor (PSD26) |
| 2943-2912 | V | | Indicates the characteristics of power state 27 | Power State 27 Descriptor (PSD27) |
| 2975-2944 | V | | Indicates the characteristics of power state 28 | Power State 28 Descriptor (PSD28) |
| 3007-2976 | V | | Indicates the characteristics of power state 29 | Power State 29 Descriptor (PSD29) |
| 3039-3008 | V | | Indicates the characteristics of power state 30 | Power State 30 Descriptor (PSD30) |
| 3071-3040 | V | | Indicates the characteristics of power state 31 | Power State 31 Descriptor (PSD31) |
| 3095-3076 | V | Varies | Shows healthy status or error code | Health indicator |

| Bytes | F = Fixed<br>V = Variable<br>X = Both | Default Value | Interpretation | Description |
|---|---|---|---|---|
| 3096 | V | Varies | Reads current negotiated PCIe link speed, as reported by PXLS register (PXCAP + 12h), bits[3:0] | Current PCIe Link Speed field (CLS) |
| 3097 | V | Varies | Reads current negotiated PCIe Link Width as reported by PXLS register (PXCAP + 12h), bits[9:4] | Negotiated Link Width (NLW) |
| 3107-3100 | V | Varies | Bootloader Version | Bootloader Version reported by the drive |
| 3109-3108 | F | 0x8086 | Company Vendor Identifier | Vendor identifier assigned by PCI-SIG (Intel) |
| 3111-3110 | F | 0x0953 | Device Identifier | Device identifier assigned by PCI-SIG (Intel) |
| 4095-3112 | V | NA | Range of bytes is allocated for vendor specific usage | Vendor Specific (VS) |

**F = Fixed**. The content of the word is fixed and does not change. For removable media devices, these values may change when media is removed or changed.
**V = Variable**. The state of at least one bit in a word is variable and may change depending on the state of the device or the commands executed by the device.
**X = F or V**. The content of the word may be fixed or variable.

**Table 32: Power State Descriptor**

| Bytes | F = Fixed<br>V = Variable<br>X = Both | Default Value | Interpretation | Description |
|---|---|---|---|---|
| 255-125 | | | | Reserved |
| 124-120 | F | 0h | Indicates the relative write latency associated with this power state | Relative Write Latency (RWL) |
| 119-117 | | | | Reserved |
| 116-112 | F | 0h | Indicates the relative write throughput associated with this power state | Relative Write Throughput (RWT) |
| 111-109 | | | | Reserved |
| 108-104 | F | 0h | Indicates the relative read latency associated with this power state | Relative Read Latency (RRL) |
| 103-101 | | | | Reserved |
| 100-96 | F | 0h | Indicates the relative read throughput associated with this power state. | Relative Read Throughput (RRT) |
| 95-64 | F | 0h | Indicates the maximum exit latency in microseconds associated with exiting this power state. | Exit Latency (EXLAT) |
| 63-32 | F | 0h | Indicates the maximum entry latency in microseconds<br>associated with entering this power state | Entry Latency (ENLAT) |
| 31-16 | | | | Reserved |
| 15-00 | F | 09C4h | 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 0.01 | Maximum Power (MP) |

**Table 33: Identify Namespace**

| Bytes | F = Fixed<br>V = Variable<br>X = Both | Default Value | Interpretation | Description |
|---|---|---|---|---|
| 7-0 | V | varies | Indicates the total size of the namespace in logical blocks. | Namespace Size (NSZE) |
| 15-8 | V | varies | Indicates the maximum number of logical blocks that may be allocated in the namespace at any point in time | Namespace Capacity (NCAP) |
| 23-16 | V | varies | Indicates the current number of logical blocks allocated in the namespace | Namespace Utilization (NUSE) |
| 24 | F | 00h | Indicates thin provisioning is not supported | Namespace Features (NSFEAT) |
| 25 | F | 06h | Defines the number of supported LBA size and metadata size combinations supported by the namespace | Number of LBA Formats (NLBAF) |
| 26 | V | 00h | Indicates metadata transferred with the extended data LBA or in separate buffer | Formatted LBA Size (FLBAS) |
| 27 | F | 03h | Indicates support for metadata transferred with the extended data LBA and in separate buffer – both are supported | Metadata Capabilities (MC) |
| 28 | V | 11h | Indicates PI supports Type 1,2,3 with PI transferred as the first 8 bytes | End-to-end Data Protection Capabilities (DPC) |
| 29 | X | 00h | Indicates type settings for the namespace | End-to-end Data Protection Type Settings (DPS) |
| 127-30 | | | | Reserved |
| 131-128 | V | MS:0, LBADS:9, RP:2 | Indicates the LBA format 0 that is supported by the controller | LBA Format 0 Support (LBAF0) |
| 135-132 | V | MS:8, LBADS:9, RP:2 | Indicates the LBA format 1 that is supported by the controller | LBA Format 1 Support (LBAF1) |
| 139-136 | V | MS:16, LBADS:9, RP:2 | Indicates the LBA format 2 that is supported by the controller | LBA Format 2 Support (LBAF2) |
| 143-140 | V | MS:0, LBADS:12, RP:0 | Indicates the LBA format 3 that is supported by the controller | LBA Format 3 Support (LBAF3) |
| 147-144 | V | MS:8, LBADS:12, RP:0 | Indicates the LBA format 4 that is supported by the controller | LBA Format 4 Support (LBAF4) |

| Bytes | F = Fixed<br>V = Variable<br>X = Both | Default Value | Interpretation | Description |
|---|---|---|---|---|
| 151-148 | V | MS:64,<br>LBADS:12, RP:0 | Indicates the LBA format 5 that is supported by the controller | LBA Format 5 Support (LBAF5) |
| 155-152 | V | MS:128,<br>LBADS:12, RP:0 | Indicates the LBA format 6 that is supported by the controller | LBA Format 6 Support (LBAF6) |
| 159-156 | | Not supported | Indicates the LBA format 7 that is supported by the controller | LBA Format 7 Support (LBAF7) |
| 163-160 | | Not supported | Indicates the LBA format 8 that is supported by the controller | LBA Format 8 Support (LBAF8) |
| 167-164 | | Not supported | Indicates the LBA format 9 that is supported by the controller | LBA Format 9 Support (LBAF9) |
| 171-168 | | Not supported | Indicates the LBA format 10 that is supported by the controller | LBA Format 10 Support (LBAF10) |
| 175-172 | | Not supported | Indicates the LBA format 11 that is supported by the controller | LBA Format 11 Support (LBAF11) |
| 179-176 | | Not supported | Indicates the LBA format 12 that is supported by the controller | LBA Format 12 Support (LBAF12) |
| 183-180 | | Not supported | Indicates the LBA format 13 that is supported by the controller | LBA Format 13 Support (LBAF13) |
| 187-184 | | Not supported | Indicates the LBA format 14 that is supported by the controller | LBA Format 14 Support (LBAF14) |
| 191-188 | | Not supported | Indicates the LBA format 15 that is supported by the controller | LBA Format 15 Support (LBAF15) |
| 383-192 | | Not supported | | Reserved |
| 4095-384 | | Not supported | Range of bytes is allocated for vendor specific usage | Vendor Specific (VS) |

**F = Fixed**. The content of the word is fixed and does not change. For removable media devices, these values may change when media is removed or changed.
**V = Variable**. The state of at least one bit in a word is variable and may change depending on the state of the device or the commands executed by the device.
**X = F or V**. The content of the word may be fixed or variable

**Table 34: LBA Format Data Structure**

| Bytes | F = Fixed<br>V = Variable<br>X = Both | Default Value | Interpretation | Description |
|---|---|---|---|---|
| 31-26 | | | | Reserved |
| 25-24 | V | Varies (2,0) | Relative Performance ranging from “best” to “degraded” | Relative Performance (RP) |
| 23-16 | V | Varies<br>(9 and 12) | Indicates the LBA data size supported. The value is reported in terms of a power of two (2^n) | LBA Data Size (LBADS) |
| 15-00 | V | Varies<br>(0, 8, 16,64, 128) | Indicates the number of metadata bytes provided per LBA based on the LBA Data Size indicated. | Metadata Size (MS) |

**F = Fixed**. The content of the word is fixed and does not change. For removable media devices, these values may change when media is removed or changed.
**V = Variable**. The state of at least one bit in a word is variable and may change depending on the state of the device or the commands executed by the device.
**X = F or V**. The content of the word may be fixed or variable.

§

# Appendix B Vital Data Structure

**Table 35: Vital Product Data Structure (VPD)**

| Address | # Bytes | Function | Programming Value | Byte | Description |
|---|---|---|---|---|---|
| 0 | 3 | Class Code | 02h | 0 | Device type and Programming Interface |
| | | | 08h | 1 | |
| | | | 01h | 2 | |
| 3 | 2 | ID | 86h | 3 | PCI-SIG Vendor ID |
| | | | 80h | 4 | |
| 5 | 20 | | Varies | 5-24 | Serial Number |
| 25 | 40 | | Varies | 25-64 | Model Number |
| 65 | 1 | PCIe Port0 Capabilities | 03h | 65 | Maximum Link Speed |
| 66 | 1 | | 08h | 66 | Maximum Link Width |
| 67 | 1 | PCIe Port1 Capabilities | 00h | 67 | Maximum Link Speed |
| 68 | 1 | | 00h | 68 | Maximum Link Width |
| 69 | 1 | Initial Power Requirements | 19h | 69 | 12V Power rail initial power requirement (W) |
| 70 | 2 | Reserved | 00h | 70-71 | |
| 72 | 1 | Maximum Power Requirements | 32h | 72 | 12V Power rail maximum power requirement (W) |
| 73 | 2 | Reserved | 00h | 73-74 | |
| 75 | 2 | Capability List Pointer | 50h | 75 | 8b address pointer to start of capability list |

**Table 36: Capability List Pointer (Out of Band Temperature Sensor)**

| Addr (Hex) | 00 | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | A | B |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | A2 | 00 | 5C | 00 | 00 | 36 | 00 | 00 | 90h | 04h | 30h | 05h |
| 5C | A2 | 00 | 00 | 00 | 00 | 32 | 00 | 00 | 90h | 04h | 30h | 05h |
| Description | Capability ID (temp) | | Next Capability (none) | | Sensor Type | SMBUS address | Reserved | | Warning Threshold | | Over Temperature | |

§

# Appendix C Out of Band Temperature Sensor Read Out

There are two temperature sensors, one for each register. Register 0x05 on address 0x1B and register 0x05 on address 0x19 contains the temperature information for the latest readout. Measured temperature is captured by bit 12 to bit 0. Data format is two's complement. Bit12 represents sign value, bit11 presents 128° C and bit0 represents 0.0625° C. Following table gives an example of the read out.

**Table 37: Register 0x05 read out format**

| Binary | Hex | Temperature |
|---|---|---|
| 1 1100 1001 0000 | 1C90 | -55° C |
| 1 1100 1110 0000 | 1CE0 | -50° C |
| 1 1110 0111 0000 | 1E70 | -25° C |
| 1 1111 1111 1111 | 1FFFF | -0.0625° C |
| 0 0000 0000 0000 | 000 | 0° C |
| 0 0000 0000 0001 | 001 | 0.0625° C |
| 0 0001 1001 0000 | 190 | 25° C |
| 0 0011 0010 0000 | 320 | 50° C |
| 0 0111 1101 0000 | 7D0 | 55° C |

§

# Appendix D SMBUS Command Response on 0x6A / 0x6B (NVMe Workgroup Defined)

**Table 38: Subsystem Management Data Structure**

| Command Code | Offset (byte) | Description |
|---|---|---|
| 0 | 00 | **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. |
| 7:0 | 01 | **Status Flags (SFLGS):** This field indicates the status of the NVM subsystem.<br>**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.<br>**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.<br>**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.<br>**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.<br>**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.<br>**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.<br>Bits 1-0 shall be set to '1'. |

<table>
<tr><th>Command Code</th><th>Offset (byte)</th><th>Description</th></tr>
<tr><td rowspan="8"></td><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.<br>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.<br>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.<br>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 the 5 seconds old.</td></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>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>
</table>

| Command Code | Offset (byte) | Description |
|---|---|---|
| | 10:09 | **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. |
| | 30:11 | **Serial Number:** 20 characters that match the serial number in the NVMe Identify Controller command response. First character is transmitted first |
| | 31 | **PEC:** An 8 bit CRC calculated over the slave address, command code, second slave addresses and returned data. Algorithm is in SMBus Specifications. |
| 32+ | 255:32 | **Vendor Specific** – This data structure shall not exceed the maximum read length of 255 specified in the SMBus version 3 specifications. Preferably length is not greater than 32 for compatibility with SMBus 2.0, additional blocks shall be on 8 byte boundaries. |

§

# *Appendix E SMBUS Command Response on 0x6A and 0x6B (Intel Specific)*

**Table 39: Command Response on 0x6A and 0x6B (Intel Specific)**

| 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 |

§

# Appendix F PCIe ID

**Table 40: PCIe ID**

| ID name | Description | Add-in Card | 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 | 0x0953 | 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 | 0x3709 | PCI Header Offset 2Ch (bits 31:16) | NA | NA |

§

# Appendix G SCSI Command Translation

The 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 the NVM Express: SCSI translation reference document at nvmexpress.org.

§

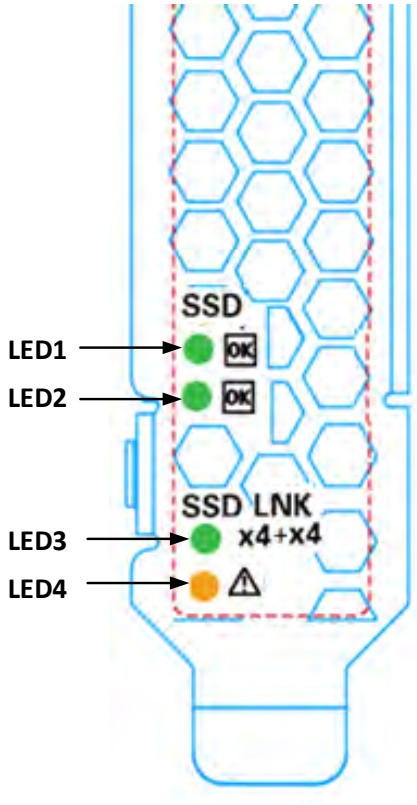
# Appendix H Add-in Card LED Decoder

On board LED are provided to indicate the status of the drive and the LEDs are defined as in the order from top to bottom on the back-plate of the device.

**Table 41: LED Functionality (Preliminary Proposal)**

| LED | Description | Blink Behavior |
|---|---|---|
| LED 1 – Activity (Green) | Shows IO activity for ASIC1 – Blinking, Controlled by Firmware | - Solid GREEN: No Read/Write Activity.<br>- Green Blinks: Read/Write Activity |
| LED 2 – Activity (Green) | Shows IO activity for ASIC2 – Blinking, Controlled by Firmware) | - Solid GREEN: No Read/Write Activity.<br>- Green Blinks: Read/Write Activity |
| LED 3 – Link Behavior | Shows Link behavior for both Physical drive controllers | - Solid GREEN: Link is established and GEN3 on all 8 lanes.<br>- OFF: Link is established, but GEN1/2 or not using all 8 lanes. |
| LED 4 – Fault (Yellow) | Fault | Solid YELLOW: Critical Warning or fault condition |

**Figure F-7-1: LED Location**




§