OCZ STORAGE SOLUTIONS | A Toshiba Group Company

Z-Drive 4500 Series

Enterprise PCI Express MLC SSD

OCZ PCIe Z-Drive 4500 SSDs are designed specifically to address enterprise storage and computing applications where demanding performance, reliability, and the total cost of ownership (TCO) are major factors.

Solution Data Sheet

OCZ Z-Drive 4500 PCIe SSD

Part Number	Description	UPC
ZD4RPFC8MT300-0800	Z-Drive 4500, PCIe, FHHL, MLC, 800GB	842024034308
ZD4RPFC8MT310-1600	Z-Drive 4500, PCIe, FHHL, MLC, 1600GB	842024034315
ZD4RPFC8MT320-3200	Z-Drive 4500, PCIe, FHHL, MLC, 3200GB	842024034322

Product Specifications

- PCI Express Gen. 2 x8 compliant
- PCIe compliant Full Height Half Length form factor
- **Processor Features**
  - Data compression, data de-duplication, and AES 128-bit encryption
  - SATA Rev. 2.6 compliant ATA feature-set
  - ATA-8 compliant
- **Capacity**
  - 800GB, 1600GB, and 3200GB
- Uses Toshiba Multi-Level Cell (MLC) NAND flash
- Uses PCIe 12V & 3.3V
- Custom configuration options allow performance aggregation across multiple cards for increased performance
- **Virtualized Controller Architecture™ (VCA) 2.0:**
  - Data path parity protection
  - Virtualizes SSD devices into a massively parallel array of memory
  - Advanced Queue Balancing Algorithm (QBA)
  - Consolidated S.M.A.R.T. support
- **Operating Systems:**
  - 32/64-bit Microsoft Windows 7, 8, 8.1
  - 32/64-bit Windows Server 2008 R2, 2012, 2013 R2
  - 64-bit Linux Red Hat Enterprise, Oracle & Centos 6.x
  - 64-bit SLES 11 SP1-3
  - 64-bit Ubuntu Server 10.04 LTS, 12.x, 13.x
  - VMware ESX/ESXi 4.1, 5.x
- **Weight**
  - 350g
- **Dimensions (L x W x H):**
  - 126.3 x 180.9 x 21.6mm
- **Compliance**
  - RoHS, FCC , CE, BSMI, C-TICK, VCCI, KCC, UL
- **Sustained Performance***
  - Max Read: Up to 2,900 MB/s
  - Max Write: Up to 2,200 MB/s
  - Random 4K Reads: Up to 252,000 IOPS
  - Random 4K Writes: Up to 44,000 IOPS
  - Random 8K Reads: Up to 155,000 IOPS
  - Random 8K Writes: Up to 31,000 IOPS
- **Power Specifications**
  - 800GB: 18.4W Idle / 20.8W Active
  - 1600GB: 20.3 Idle / 23.1W Active
  - 3200GB: 20W Idle / 22.8W Active
- **Temperature**
  - Operating: 0°C to 55°C
  - Non-Operating: -45°C to 85°C
- Temperature sensing and thermal throttling without sacrificing performance under normal operation
- **Airflow Requirements**
  - 300 LFM (Linear Feet per Minute) at 25°C
  - 550LFM (Linear Feet per Minute) at 40°C
- **Reliability**
  - Mean Time Between Failures (MTBF): 2 million hours
  - Endurance Rating: Minimum of 0.68PB (800GB), 1.3PB (1600GB), 2.5PB (3200GB) over 5 year period
  - Unrecoverable Bit Error Rate (UBER): Less than 1 sector per 10^{17} bits read
  - RAID technology to provide data recovery in case of block failure
  - Field-upgradable firmware

* ZD4RPFC8MT320-3200 performance

Table of Contents

1. Introduction

The OCZ Z-Drive 4500 PCIe SSD Series are designed specifically to address the demanding performance and reliability requirements of today's enterprise-class storage and computing applications. The series leverages 19 nanometer (nm) Multi-Level Cell (MLC) NAND flash covering 800GB, 1.6TB and 3.2TB usable capacities. The series delivers even higher performance in comparison to OCZ's Z-Drive R4 Series and features a more robust architectural design with a priority on data reliability and drive endurance. The available Z-Drive 4500 models include the Model ZD4RPFC8MT300-0800 (supporting 800GB capacity), Model ZD4RPFC8MT310-1600 (supporting 1.6TB capacity) and Model ZD4RPFC8MT320-3200 (supporting 3.2TB capacity).

Product Overview:

The Z-Drive 4500 PCIe SSD Series are driven by OCZ's proprietary Virtualized Controller Architecture™ (VCA) that dynamically reorders the storage commands and processes them using eight (8) LSI® SF-2582 processors. By utilizing the full processing bandwidth of eight enterprise-class processors, the storage system runs more efficiently while delivering RAID-like performance.

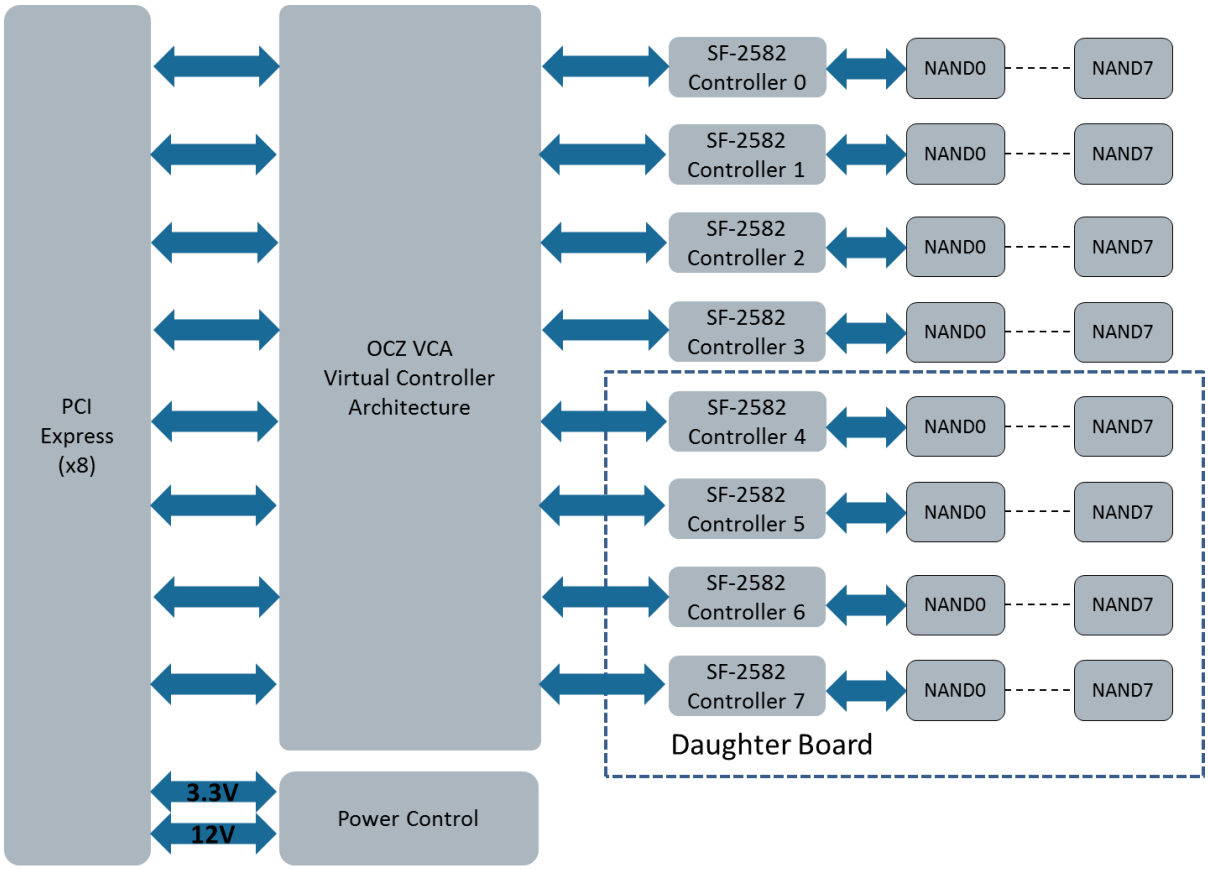
Based on SNIA test methodologies, the Z-Drive 4500 Series delivers industry-leading sustained performance for MLC-based PCIe edge cards including maximum read bandwidth of up to 2,900 MB/s, maximum write bandwidth of up to 2,200 MB/s, random 4K block read throughput of up to 252,000 input/output operations per second (IOPS) and random 4K block write throughput of up to 76,000 IOPS. As a result, Z-Drive 4500 models are ideally suited for I/O read and write intensive enterprise applications where high storage capacities coupled with low power NAND flash results in higher bandwidth and IOPS performance per watt.

With a priority on data reliability and drive endurance, the Z-Drive 4500 Series features an advanced suite of tools and capabilities including both power loss data protection (that protects in-flight write operations in the event of a power failure) and end-to-end data path protection (that performs data integrity checks at every data juncture ensuring that corrupted data is detected and not propagated). The Z-Drive 4500 PCIe SSDs have an excellent 2.2 million hour mean time before failure (MTBF) ensuring reliability over long term usage and backed by a 5-year warranty.

Each model within the Z-Drive 4500 Series is supported by OCZ's new Windows Accelerator (WXL) Software -- a flash management and caching solution for Microsoft Windows Server applications that enables IT managers to deliver low-latency flash deployable as a local flash volume, a flash cache for HDD volumes or as a combination of both. When deployed for caching, WXL Software performs statistical 'out-of-band' processing that intelligently caches the most frequently accessed data on Z-Drive 4500 flash. A cache warm-up and analysis mechanism is also featured enabling important and demanding analytical jobs to be loaded onto the Z-Drive 4500's flash cache in advance to assure that this critical data is available to the application at the exact time it is needed.

For flash-based virtualization, the Z-Drive 4500 Series works in conjunction with OCZ's VXL Virtualization Software distributing flash cache resources on-demand across virtual machines (VMs) to accelerate application performance. It distributes the flash between VMs based on need making sure that no VM inefficiently occupies flash when it could be better used elsewhere in the environment. As a result, the Z-Drive 4500's flash cache is optimally utilized at all times regardless of how many VMs are running concurrently, data traffic to and from the SAN is reduced, and critical data is locally available in the Z-Drive 4500 card for immediate use by VMs.

Table 1.0



2. PCIe x8 Connector Pin Out

The PCIe x8 signal names and pinout are listed in the table below.

Table 2.0 – PCIe x 8 Connector Pinout

Pin	Side B Connector		Side A Connector	
#	Name	Description	Name	Description
1	+12V	+12V volt power	PRSNT#1	Hot plug presence detect
2	+12V	+12V volt power	+12V	+12V volt power
3	+12V	+12V volt power	+12V	+12V volt power
4	GND	Ground	GND	Ground
5	SMCLK	SMBus clock	JTAG2	TCK
6	SMDAT	SMBus data	JTAG3	TDI
7	GND	Ground	JTAG4	TDO
8	+3.3V	+3.3 volt power	JTAG5	TMS
9	JTAG1	+TRST#	+3.3V	+3.3 volt power
10	3.3Vaux	3.3 volt power	+3.3V	+3.3 volt power
11	WAKE#	Link reactivation	PWRGN	Power Good
Mechanical Key				
12	RSVD	Reserved	GND	Ground
13	GND	Ground	REFCLK+	Reference Clock, Differential pair
14	HSOp(0)	Transmitter Lane 0, Differential pair	REFCLK-	
15	HSOn (0)		GND	Ground
16	GND	Ground	HSIp(0)	Receiver Lane 0, Differential pair
17	PRSNT#2	Hot plug detect	HSIn(0)	
18	GND	Ground	GND	Ground
19	HSOp(1)	Transmitter Lane 1, Differential pair	RSVD	Reserved
20	HSOn(1)		GND	Ground
21	GND	Ground	HSIp(1)	Receiver Lane 1, Differential pair
22	GND	Ground	HSIn(1)	
23	HSOp(2)	Transmitter Lane 2, Differential pair	GND	Ground
24	HSOn(2)		GND	Ground
25	GND	Ground	HSIp(2)	Receiver Lane 2, Differential pair
26	GND	Ground	HSIn(2)	
27	HSOp(3)	Transmitter Lane 3, Differential pair	GND	Ground
28	HSOn(3)		GND	Ground
29	GND	Ground	HSIp(3)	Receiver Lane 3, Differential pair
30	RSVD	Reserved	HSIn(3)	
31	PRSNT#2	Hot plug detect	GND	Ground
32	GND	Ground	RSVD	Reserved
33	HSOp(4)	Transmitter Lane 4, Differential pair	RSVD	Reserved
34	HSOn(4)		GND	Ground
35	GND	Ground	HSIp(4)	Receiver Lane 4, Differential pair
36	GND	Ground	HSIn(4)	
37	HSOp(5)	Transmitter Lane 5,	GND	Ground

Pin	Side B Connector		Side A Connector	
38	HSOn(5)	Differential pair	GND	Ground
39	GND	Ground	HSIp(5)	Receiver Lane 5, Differential pair
40	GND	Ground	HSIn(5)	
41	HSOp(6)	Transmitter Lane 6, Differential pair	GND	Ground
42	HSOn(6)		GND	Ground
43	GND	Ground	HSIp(6)	Receiver Lane 6, Differential pair
44	GND	Ground	HSIn(6)	
45	HSOp(7)	Transmitter Lane 7, Differential pair	GND	Ground
46	HSOn(7)		GND	Ground
47	GND	Ground	HSIp(7)	Receiver Lane 7, Differential pair
48	PRSNT#2	Hot plug detect	HSIn(7)	
49	GND	Ground	GND	Ground

3. Certification & Compliance

Table 3.0 – Certification

Certification/Compliance	Description
CE Compliant	Indicates conformity with the essential health and safety requirements set out in European Directives Low Voltage Directive and EMC directive.
FCC Compliant B	FCC Class B devices are those that are for use in a commercial, industrial, or business environment. Class B devices are those that are marketed for use in the home.
RoHS Compliant	Restriction of Hazardous Substance Directive
BSMI	Bureau of Standards, Metrology and Inspection
C-TICK	Australia's equivalent to the FCC Declaration of Conformity
VCCI	Voluntary Control Council for Interference by Information Technology Equipment or VCCI is the Japanese body governing RF emissions
KCC	Korea Communications Commission equivalent to the FCC Declaration of Conformity
UL*	UL safety consulting and certification company
VMware driver certification*	I/O Vendor Partner Program (IOVP) certification

* In Process

4. Product Specifications

4.1. Capacity

Table 4.0 – Capacity Description

RAW Capacity	Usable Capacity	Total User Addressable Sectors in LBA Mode
1TB	800GB	745.28GB (LBA48 = 195371568)
2TB	1,600GB	1490.48GB (LBA48 = 390721968)
4TB	3,200GB	2980.77GB (LBA48 = 781422768)

Notes:

- Partitioned and formatted capacity in Microsoft Windows
- 1 sector = 512 byte
- 1GB = 1 billion bytes, actual formatted capacity less

Performance - MB/s Throughput

Table 4.1 Maximum Sustained Sequential Read, Write Bandwidth (MB/s)

Metric	800GB	1600GB	3200GB
Max Sequential Read (MB/s)[1]	2,900	2,900	2,900
Max Sequential Write (MB/s)[1]	1,300	1,900	2,200

[1]Based on 100% Incompressible Data, block size 1024KB: 4-Threads @ QD64 per thread

Table 4.2 - Maximum Sustained 4K Random IOPS (Input/Output Operations per Second)

Metric	800GB	1600GB	3200GB
4KB Random Read[1]	164,000	215,000	252,000
4KB Random Write[1]	57,000	76,000	43,000

[1]Based on SNIA methodology and workloads, see appendix for details

Table 4.3 - Maximum Sustained 8K Random IOPS

Metric	800GB	1600GB	3200GB
8KB Random Read[1]	120,000	144,000	155,000
8KB Random Write[1]	27,000	33,000	31,000

[1]Based on SNIA methodology and workloads, see appendix for details

4.2. Endurance Rating

Endurance rating of the SSD is represented in TBW, terabyte(s) written by a host to the SSD

Table 4.4 – TBW Rating

Description	800GB	1600GB	3200GB
Minimum TBW product lifetime[1]	680	1300	2500

[1]Based on incompressible data

4.3. Electrical Characteristics

4.3.1. Supply Voltage

Table 4.5 – Operating Voltage

Description	Min	Max	Unit
Operating Voltage for 3.3 V (+/- 9%)	3.003	3.597	V
Operating Voltage for 12 V (+/- 8%)	11.04	12.96	V

4.3.2.Power Consumption

Table 4.6 – Power Consumption

Mode	800GB	1,600GB	3,200GB	Unit
Active	18.4	20.3	20.0	W
Idle	20.8	23.1	22.8	W

4.4. Environmental Conditions

4.4.1.Temperature

Table 4.7 - Temperature

Mode	Min	Max	Unit
Operating	0	55[1]	°C
Non-Operating	-45	85	°C

Please note: Adequate airflow most be provided at higher temperatures; airflow requirements: 300 LFM (Linear Feet per Minute) at 25°C and 550 LFM (Linear Feet per Minute) at 40°C are recommended minimums

4.5. Reliability

Table 4.8 - Reliability

Parameter	Value
MTBF	2,200,000 Hours
MLC Program Erase Cycles (P/E)	3,000 Cycles
ECC (BCH)	55 its per 512 bytes of data

4.5.1.Mean Time Between Failures

The Mean Time Between Failures (MTBF) is calculated based on a Part Stress Analysis. It assumes nominal voltage, with all other parameters within specified range. This specification is based on the MTBF rating of the LSI SF-2582 controller.

4.5.2.Endurance Rating

The endurance rating specifies the number of terabytes that may be written to the SSD over a period of time. Write amplification and program / erase cycles can significantly impact the endurance rating.

4.5.3. Program / Erase Cycles:

Specifies the maximum number of cycles the NAND flash chips can be programmed and erased before losing their ability to hold a charge.

4.5.4. ECC

Error-Correcting Code defines the number of correctable symbols for a given size of data. The LSI SF-2582 processor is capable of correcting up to 512 bits per 512 bytes using a BCH algorithm.

4.6. Additional Features

4.6.1. Compression and De-Duplication

Drives use both compression and de-duplication to improve performance and prolong drive life.

4.6.2. Encryption

Z-Drive 4500 SSDs use AES 128-bit encryption on all meta-data and user data. User key is not enabled by default, but data is encrypted at all times for Self Encrypting Device (SED).

4.6.3. Proprietary Virtualized Controller Architecture 2.0 (VCA)

This technology enables the Z-Drive 4500 to appear to the host as a single device and:

- Provides a multifaceted virtualization layer that interfaces with the host system
- Enables TRIM command support, firmware updates, secure erase, and SMART health monitoring
- Extends NAND flash life at a block level
- Manages redundancy
- Features internal storage controller delivering highly efficient performance aggregation while reducing the burden on host resources

4.6.4. Data Redundancy

Internal SSD RIAD functionality enables the Z-Drive 4500 to recover data from a failed NAND flash block (per NAND controller) by rebuilding the data and relocating it to the known good flash array. This data recovery process is transparent to the host.

5. ATA/SATA Power management

The LSI SF-2582 supports ATA and SATA power management modes as described below:

5.1. ATA Power Modes

The ATA power modes supported by the LSI SF-2582 controller are:

- ACTIVE
- IDLE
- STANDBY
- SLEEP

5.2. SATA Link Power States

The SATA power states supported by the LSI SF-2582 controller are:

- ACTIVE: PHY Ready, full power, Tx & Rx operational
- PARTIAL: Reduced power, resumes in under 10usec
- SLUMBER: Reduced power, resumes in under 10msec
- HIPM: Host-Initiated Power Management
- DIPM: Device-Initiated Power Management
- AUTO-SLUMBER: Automatically transitions to partial to slumber

6. S.M.A.R.T. Attribute Support

This section describes the S.M.A.R.T. attributes supported by the LSI SF-2582 processor

Table 7.0 - Attribute ID Numbers

ID	Attribute Name	Description
1	Raw Read Error Rate	Raw error rate related to ECC errors. Correctable and uncorrectable RAISE errors are included in the error event count. (UECC + URAISE)
5	Retired Block Count	Indicates the total number of retired blocks, showing both manufacturing bad blocks and growing bad blocks. As the blocks are retired, the normalized value decreases to 0 from 100
9	Power On Hours (POH)	Counts of hours in power-on state. The raw value of this attribute shows total count of hours in power-on state. The value is updated hourly
12	Device Power Cycle Count	Indicates the count of drive power on/off cycles since the manufactured date
13	Soft Read Error Rate	Indicates read errors that are corrected by the ECC on the fly
100	Gigabytes Erased	Counts the number of Flash gigabytes erased across the entire drive over the life of the drive. The incremental resolution is 64GB
170	Reserve Block Count	Returns current number of reserve (over-provisioned) blocks remaining. This attribute is related to attribute 5 (Retired Block Count)
171	Program Fail Block Count	Counts the number of flash program failures since the drive was deployed
172	Erase Fail Block Count	Counts the number of flash erase failures since the drive was deployed
174	Unexpected Power Loss Count	Counts the number of unexpected power loss events without receiving a Sleep command or Standby Immediate since the drive was deployed.
177	Wear Range Delta	Returns the percent difference in wear between the most-worn block and least-worn block.
181	Program Fail Count	4 bytes used to show the number of program failures since the drive was deployed. It is identical to attribute 171
182	Erase Fail Count	4 bytes used to show the number of block erase failures since the drive was deployed. It is Identical to attribute 172
184	Reported I/O Error Detection Code Errors	Tracks the number of end-to-end CRC errors encountered during host initiated reads and writes. The count is based on the number of errors since the last power applied to the drive
187	Reported Uncorrectable Errors	This attribute tracks the number of uncorrectable RAISE (URAISE) errors reported back to the host for all data access commands. It is identical to attribute 198
194	Temperature	Temperature assuming an on-board sensor connected via ISTW interface to the SSD processor
195	On the fly ECC Uncorrectable Error Count	Tracks the number of uncorrectable errors (UECC) by ECC on the fly but correctable by the RAISE. It is identical to the attribute 201
196	Reallocation Event Count	Tracks the number of blocks that fail programming which are reallocated as a result. These failed bad blocks are considered as grown bad blocks
198	Uncorrectable Sector Count	Tracks number of uncorrectable RAISE errors when reading/writing a sector since the drive was deployed. It's identical to the attribute 187.

ID	Attribute Name	Description
199	SATA R Read Errors (CRC) Count	Tracks SATA transmit and receive R-Read errors since the last power applied to the drive
201	Uncorrectable Soft Read Error Rate	Tracks number of soft read errors that cannot be fixed by ECC on-the-fly and requires deep recovery via RAISE. (i.e. UECC). It's identical to the attribute 204 and attribute 195
204	Soft ECC Correction Rate	Tracks number of errors corrected by RAISE that cannot be fixed by ECC on-the-fly and requires ECC (multilevel) to correct. (i.e. UECC). It's identical to attribute 201 and attribute 195
230	Life Curve Status	A life curve used to help predict life in terms of the endurance based on the number of writes to flash.
231	SSD Life Left	Measures of the estimated life left, based on a combination of PE cycles and available reserve blocks. 100 is a new drive, 10=replace as it has sufficient reserved blocks but the PE cycles have been used, 0=insufficient reserved blocks, drive is read only to allow recovery of data on the drive.
232	Available Reserved Space	The number of reserved blocks remaining. This is stored in 4 bytes. It is the same as attribute 170, but expressed in gigabytes.
241	Lifetime Writes from Host	Indicates the total amount of data written from hosts since the drive was deployed. This is stored in 4 bytes. The number stored represents the number of bytes written by the host to the drive, in 64GB increments.
242	Lifetime Reads from Host	Indicates the total amount of data read to hosts since the drive was deployed. This is stored in 4 bytes. The number stored represents the number of bytes read by the host to the drive, in 64GB increments.

7. AC, DC and Timing Specifications

This section describes "Virtualized Controller Architecture" DC characteristics. Stresses beyond those listed in the following table can damage the device. These are stress ratings only. Functional operation of the device or beyond these values is not implied.

The Z-Drive 4500 timings conform to the PCIe Specification, Rev 2.0 and the Serial Attached SCSI-2 (SAS-2).

Table 8.0 – Absolute Maximum Stress Ratings

Symbol	Parameter	Min	Typ	Max	Unit
AVDD[8:0]	Absolute Analog Power for PCIe PHY	1.62	1.8	1.98	V
VAA[7:0], VAA_ANA	Absolute Analog Power for SATA PHY, Chip PLL	2.25	2.5	2.75	V
VDD	Absolute Power for Digital Core	0.9	1.0	1.1	V
VDD01/VDD02	Absolute Digital I/O Power	3	3.3	3.6	V

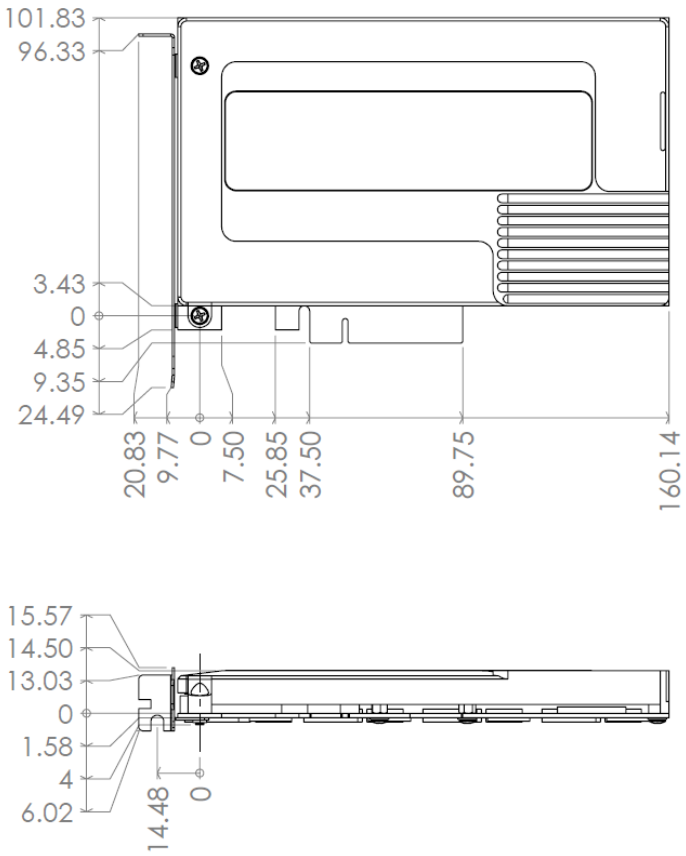
Table 8.1 – Recommended Operating Conditions

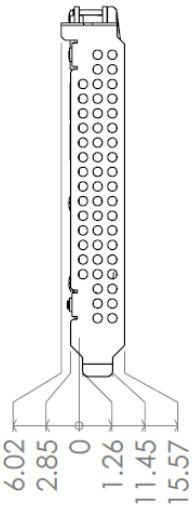
Symbol	Parameter	Min	Typ	Max	Unit
AVDD[8:0]	Analog Power for PCIe PHY	1.71	1.8	1.98	V
VAA[7:0], VAA_ANA	Analog Power for SATA PHY, Chip PLL	2.38	2.5	2.63	V
VDD	Digital Core Power	0.95	1.0	1.05	V
VDD01/VDD02	Digital I/O Power	3.14	3.3	3.47	V
ISET, PIN_ISET	Internal Bias Reference	5.74	6.04	6.34	KΩ
T_A	Ambient Operating Temperature	0	-	70	°C
T_J	Junction Operating Temperature	0	-	125	°C

Table 8.2 – DC Electrical Characteristics

Symbol	Parameter	Min	Typ	Max	Unit
I_{VAA}	Analog power for PCIe PHY 1.8v	-	0.78	-	A
I_{VAA}	Analog Power for SAS/SATA PHY 2.5V, Chip PLL	-	0.78	-	A
I_{VDD}	Digital Core Power	-	2.0	-	A
I_{VDDO}	Digital I/O Power	-	50	-	mA
I_{IL}	Input Low Voltage of Digital I/O	-0.4	-	0.8	V
I_{IH}	Input High Voltage of Digital I/O	2.0	-	3.6	V
I_{OL}	Output Low Voltage of Digital I/O	-	0.13	-	V
I_{OH}	Output High Voltage of Digital I/O	2.0	VDDO1/VDDO2	-	V

8. Mechanical Specifications





Notes:

- All dimensions are in millimeters

9. Glossary

Term	Definition
ATA	Advanced Technology Attachment
ATAPI	Advanced Technology Attachment Packet Interface
DMA	Direct Memory Access
ECC	Error-Correcting Code
EXT	Extended
GB	
GC	
HDD	Hard Disk Drive
Hot Plug	A term used to describe the removal or insertion of a SATA hard drive when the system is powered on.
IOPS	Input output operations per second
LBA	Logical Block Address
MB	Mega-byte defined as $1x10^6$ bytes
MTBF	Mean time between failures
NCQ	Native Command Queuing. The ability of the SATA hard drive to queue and re-order commands to maximize execution efficiency.
NOP	No Operation
OS	Operating System
Port	The point at which a SATA drive physically connects to the Processor.
P/E	Program / Erase cycles, defines NAND lifecycle
SAS	Serial Attached SCSI
SATA	Serial ATA
SFF	Small Form Factor
SMART	Self-Monitoring, Analysis and Reporting Technology: an open standard for deping hard drives and software systems that automatically monitors a hard drive's health and reports potential problems.
SSD	Solid State Drive
TRIM	ATA8 Command informing the drive when sectors no longer contain valid data

10.Appendix:

10.1. SNIA Workload

Workload Profile	SSD State	Block Size	Access Pattern	Data Pattern	Starting Sector	Maximum Size	Alignment	Threads	QD per Thread
4K Random	Steady State	4KB	Random	Incompressible	0 (RAW)	SSD Max	4K	4	128
8K Random	Steady State	8KB	Random	Incompressible	0 (RAW)	SSD Max	4K	4	128