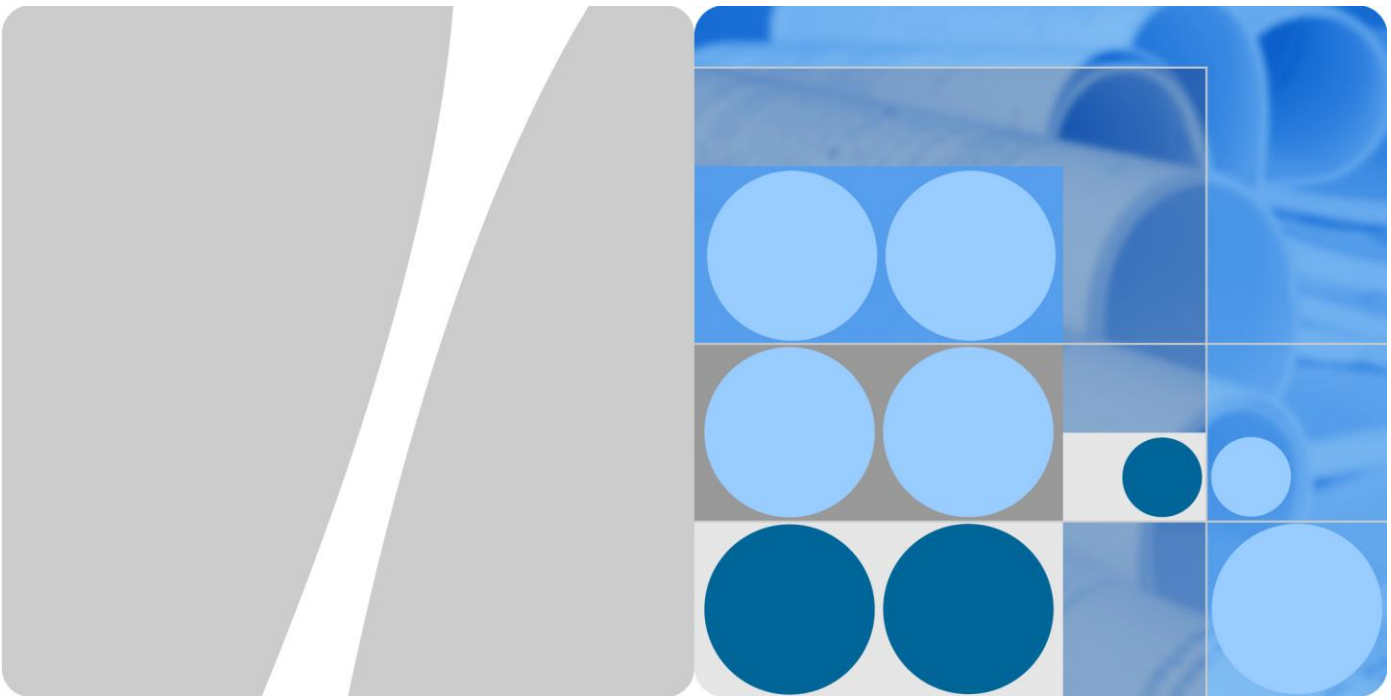


Huawei ES3000 V5 NVMe PCIe SSD White Paper

Huawei ES3000 V5 NVMe PCIe SSD White Paper

Issue 01

Date 2018-10-31

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions


and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: http://www.huawei.com

Email: support@huawei.com

About This Document

Overview

This document describes the system design, features, and specifications of Huawei ES3000 V5 NVMe SSDs.

Intended Audience

This document is intended for:

- Technical support engineers
- Maintenance engineers
- Customers

Symbol Conventions

The symbols that may be found in this document are defined as follows:

Symbol	Description
DANGER	Indicates an imminently hazardous situation which, if not avoided, will result in death or serious injury.
WARNING	Indicates a potentially hazardous situation which, if not avoided, could result in death or serious injury.
CAUTION	Indicates a potentially hazardous situation which, if not avoided, may result in minor or moderate injury.
NOTICE	Indicates a potentially hazardous situation which, if not avoided, could result in equipment damage, data loss, performance deterioration, or unanticipated results. NOTICE is used to address practices not related to personal injury.
NOTE	Calls attention to important information, best practices and tips. NOTE is used to address information not related to personal injury, equipment damage, and environment deterioration.

Change History

Issue	Date	Description
02	2018-08-21	Modified the Technical Specifications section.
01	2018-07-06	This issue is the first official release.

Contents

1 Introduction

1.1 Overview

Huawei ES3000 V5 NVMe PCIe SSD (ES3000 V5 for short) is an enterprise-level storage acceleration product that features high performance, quick responses, and high reliability, greatly improving storage I/O performance. Compatible with mainstream operating systems and virtualization systems, the ES3000 V5 helps improve service performance of databases, virtualization, high-performance computing (HPC), searching, and other applications, reducing customers' total cost of ownership (TCO).

The ES3000 V5 uses the PCIe 3.0 x4 interface and consists of two series:

- ES3500P series: standard 2.5-inch disk with the SFF-8639 connector
  Designed for scenarios with intensive-read workloads.
- ES3600P series: standard 2.5-inch disk with the SFF-8639 connector
  Designed for scenarios with mixed read/write workloads.
- ES3600C V5 series: standard half-height half-length (HHHL) PCIe card
  Designed for scenarios with mixed read/write workloads.

The ES3000 V5 comes in various single-disk capacities and specifications to meet requirements of different applications.

- ES3500P series single-disk capacities: 1000 GB, 2000 GB, 4000 GB, and 8000 GB
- ES3600P series single-disk capacities: 800 GB, 1600 GB, 3200 GB, and 6400 GB
- ES3600C series single-disk capacities: 800 GB, 1600 GB, 3200 GB, and 6400 GB

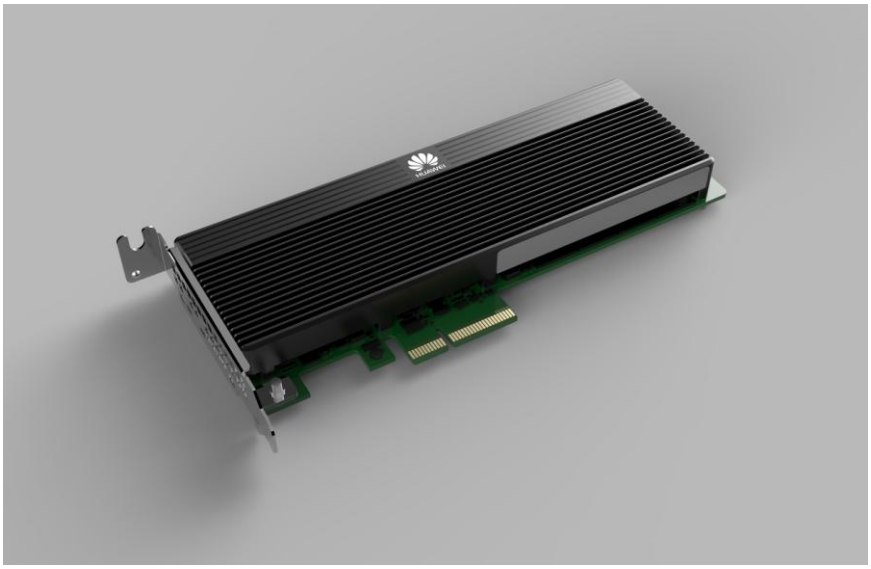
The following figures show the ES3000 V5.

Huawei ES3000 V5 NVMe PCIe SSD White Paper

Figure 1-1 ES3500P V5

Figure 1-2 ES3600P V5

Figure 1-3 ES3600C V5



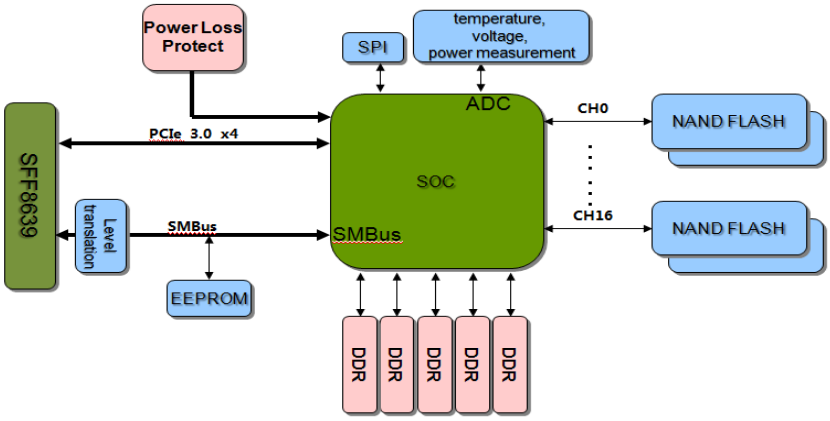
1.2 System Design

The ES3000 V5 has the following system design features:

- Uses Huawei's Hi1812E SSD controller. The SSD architecture design fully realizes the potential of the SSD controller. Hi1812E is an ASIC controller based on Huawei's SSD flash translation layer (FTL) algorithm and low-density parity-check (LDPC) error correcting algorithm, supporting the NVMe 1.3 specification and featuring high performance and low power consumption.
- Provides innovative NVMControl technical sets based on Hi1812E to enable high performance and reliability of the SSD.
- Provides enhanced device management functions. The standard NVMe interface eliminates the need for drivers built in OSs and allows the SSD to serve as an OS disk. Comprehensive hot swap functions realize easy maintenance. Out-of-band management functions comply with the NVMe specification and work with Huawei servers to enable intelligent monitoring of device health status.

Figure 1-4 shows the ES3000 V5 system architecture.

Figure 1-4 ES3000 V5 system architecture



2 Features

2.1 Performance Features

- Uses Huawei Hi1812E ASIC SSD controller.
- Supports the NVMe specification and uses multi-queue I/O technology to improve the SSD performance and quality of service (QoS).
- Adopts the innovative NVMControl technology to ensure high performance and reliability.

2.2 Reliability Features

- Uses the enhanced LDPC algorithm to provide higher error correction capability than that required by flash memory chips, ensuring device reliability.
- Restores data after a media data occurs by using the embedded intelligent RAID-like algorithm.
- Adopts the patented flexible RAID algorithm to provide RAID protection for recovered data after channels fail.
- Uses the intelligent wear leveling algorithm to intelligently level the flash memory wear and improve endurance.
- Supports the advanced flash memory access technology to apply the read retry and adaptive read technologies of flash memory chips and ensure data validity.
- Supports intelligent FSP algorithms to provide faster and more reliable data storage services by working with the 3D TLC feature.
- Adopts the data inspection technology to periodically inspect data and prevent errors.
- Supports power failure protection to ensure that data on the SSD is not lost when a power failure occurs on the server.
- Provides end-to-end data protection through the DIF domain, and data check for every data transmission between modules within the SSD.

2.3 Maintenance Features

- Supports in-band online upgrades to facilitate routine maintenance.
- Provides device information in a centralized manner, including the model, capacities, temperature, remaining service life, and health status.
- Provides abundant command-line interface (CLI) tools to facilitate device daily maintenance.
- Provides the manufacture dates and serial numbers to facilitate asset management.

3 Appearance and Mechanical

3.1 Appearance

The ES3000 V5 is a standard 2.5-inch disk. It can be installed in universal servers and storage array.

3.1.1 2.5-inch SSD in the Disk Tray

Figure 3-1 Front view of a 2.5-inch disk in a hard disk tray

The ES3000 V5 supports operating status indicators provided by disk enclosures. The amber/blue indicator in 3.1.1 2.5-inch SSD in the Disk Tray is the Fault indicator, and the green indicator is the Activity indicator. The color of the indicator is closely related to the design of the disk enclosures. When the ES3000 V5 is used in Huawei servers, the status indicators are defined as follows:

Table 3-1 Status indicators definition

Activity Indicator (Green)	Fault Indicator (Amber/Blue)	Status
Off	Off	• The SSD is powered off.
On	Off	• The SSD is operating properly.
Blinking (2 Hz)	Off	• Data is being read or written.

Activity Indicator (Green)	Fault Indicator (Amber/Blue)	Status
Off	Blinking (2 Hz)	• The SSD is located or the OS is ejecting the SSD for hot swap.
Off	Blinking (0.5 Hz)	• The SSD is ejected by the OS and can be hot removed.
Off/On	On	• The SSD is faulty.

3.1.2 SSD Card Front View

Figure 3-2 ES3600C card front view

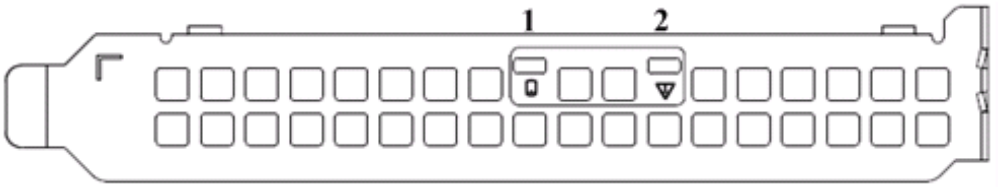


Table 3-2 Indicator description

1	Active indicator
2	Fault indicator

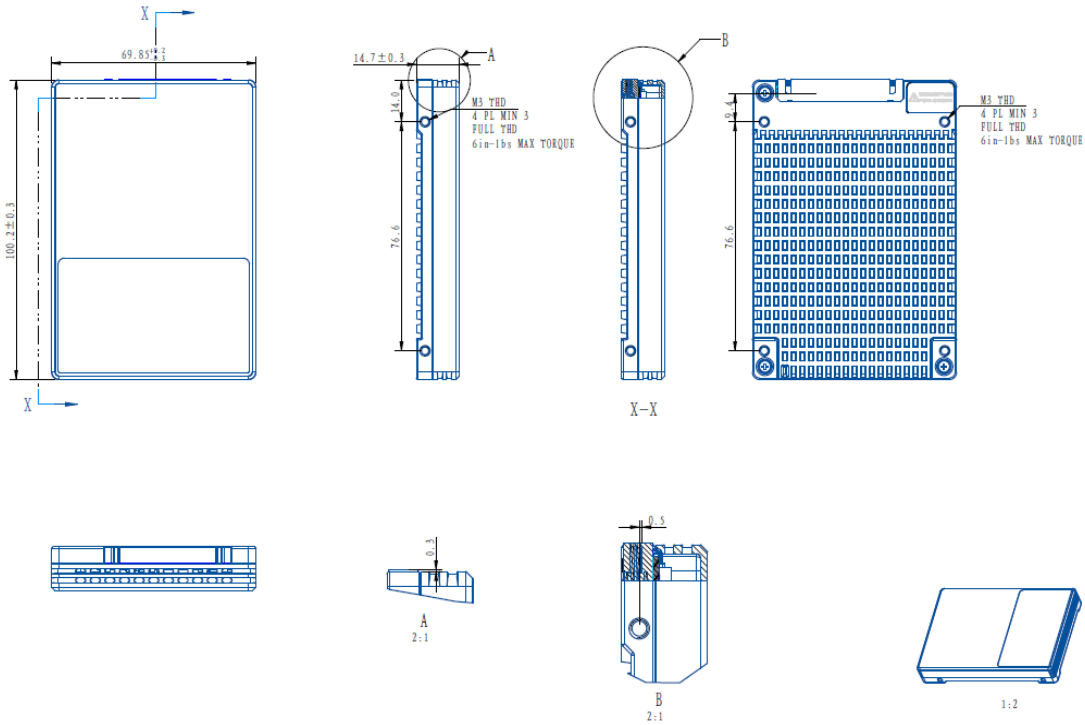
Table 3-3 Definitions of SSD card status indicators

Active Indicator (Green)	Fault Indicator (Yellow)	Status
Off	Off	• The PCIe card is not powered on.
Off	On	• Firmware of the PCIe card fails to be loaded.
Off	Blinking (0.5 Hz)	• The PCIe card is being initialized.
On	Off	• The NVMe disk can be detected.
Blinking (2 Hz)	Off	• The NVMe disk is being read and written.
On	On	• The device is faulty.

3.2 Mechanical

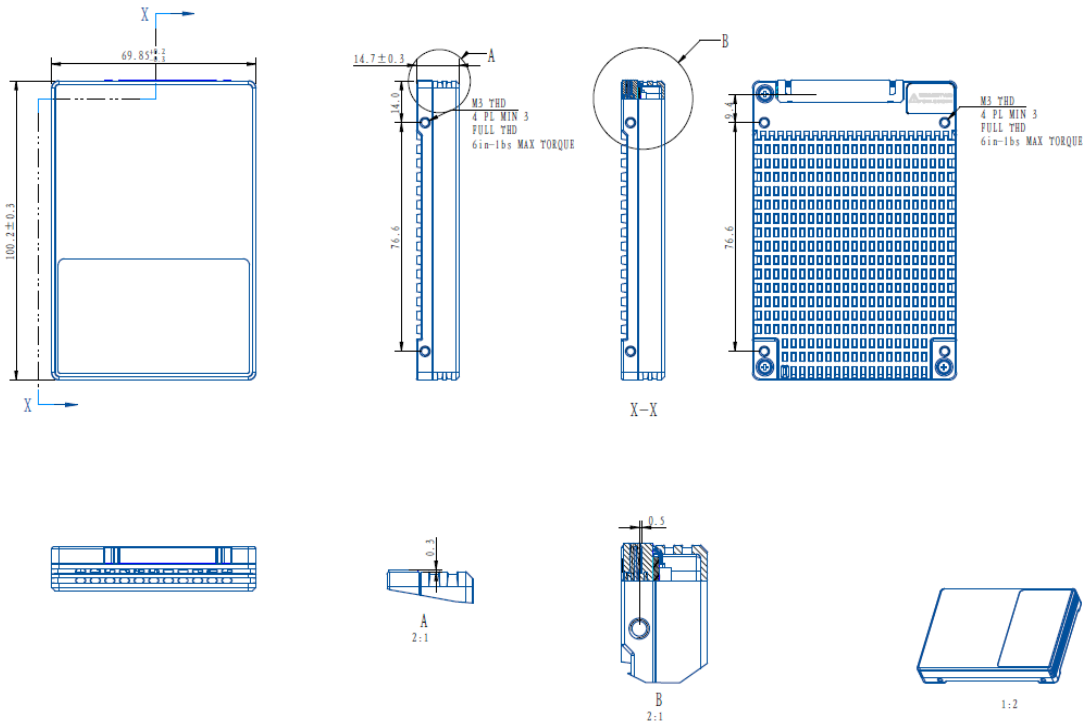
Note: All dimensions are in millimeters.

3.2.1 2.5-inch SSD



X – Length	Y – Width	Z – Height
100.2+/–0.3	69.85+0.2/–0.3	14.7+/–0.3

3.2.2 2.5-inch SSD1



X – Length	Y – Width	Z – Height
100.2+/–0.3	69.85+0.2/–0.3	14.7+/–0.3

4 Technical Specifications

4.1 Basic Specifications

4.1.1 ES3500P V5

Table 4-1 ES3500P V5 specifications[1] (LBA[2] format: 512 B)

Feature	Specifications			
Model	HWE52P431T0 L002N	HWE52P432T0 L002N	HWE52P434T0 L002N	HWE52P438T0 L002N
Product name	ES3500P V5-1000	ES3500P V5-2000	ES3500P V5-4000	ES3500P V5-8000
Form	2.5-inch disk PCIe 3.0	2.5-inch disk PCIe 3.0	2.5-inch disk PCIe 3.0	2.5-inch disk PCIe 3.0
Capacity	1000 GB[3]	2000 GB	4000 GB	8000 GB
Flash chip type	3D TLC	3D TLC	3D TLC	3D TLC
Sequential read bandwidth (MB/s)[4]	3,500	3,500	3,500	3,500

Feature	Specifications			
4KB random read IOPS (Steady)[5]	420,000	740,000	825,000	825,000
Average read latency@1QD (μs)[6]	96	96	96	88
Sequential write bandwidth (MB/s)	1,000	1,900	3,200	2,600
4KB random write IOPS (Steady/Max.)[7]	63,000	90,000	110,000	150,000
Average write latency@1QD (μs)	17	15	15	14
Hybrid (R/W: 7/3) random 4KB IOPS (Steady)	130,000	190,000	265,000	330,000
PBW (lifetime random PB written)[8]	1.825 PB	3.65 PB	7.3 PB	14.6 PB
DWPD (disk writes per day)[9](JESD219)	1	1	1	1
Average power consumption (idle, activity)[10]	5 W, 13 W	5 W, 13.8 W	5.5 W, 18.8 W	5.7 W, 21 W
Weight (g)	182			
Power-off protection	Supported			
Chip failure protection	Supported			
Trim	Supported			
Mean time between failures (MTBF)	2.5 million hours			
Annual failure	≤ 0.35%			

Feature	Specifications
rate (AFR)	
Uncorrectable bit error rate (UBER)	10^{-18}
Data retention (power off)[11]	3 months at 40 ℃ (104 ℉)
Data retention of a new SSD used a spare (power off)[12]	12 months at 40℃ (104 ℉)

NOTE

1. The performance specifications are test at the maximum power.
2. Logical block addressing (LBA) is a common mechanism for describing data blocks on storage devices. It is usually used on hard disks. LBA is a special simple linear addressing mechanism, and an address block is indexed by an integer. For example, LBA 0 indicates the first block, LBA 1 indicates the second block, and the list goes on.
3. GB = 1,000,000,000 bytes.
4. MB/s = 1,000,000 bytes per second. Performance measured using FIO on CentOS 7.0 with 1 MB (1,048,510 bytes) transfer size 512 Queue Depth on a sequential workload and number of threads equal to 1. Sequential write bandwidth is likeness.
5. Performance measured using FIO on RHEL 7.0 with 4 KB transfer size 64 Queue Depth 8 threads on random read in steady state.
6. μs = microsecond. Average read latency measured using 1 thread 1 QD 4 KB transfer size on a random workload. Average write latency is likeness.
7. Performance measured using FIO on RHEL 7.0 with 4 KB transfer size 64 Queue Depth 8 threads on random write in steady state.
8. 1PB = 10^{15} bytes. PBW measured using 4KB transfer size with 4KB aligned LBA on a random workload.
9. The DWPD is tested according to the JESD219 standard. An SSD can be used for five years if the the DWDP stays below the specified value. Otherwise, the SDD service life will be affected.
10. Activity power consumption measured using sequential writes.

    ES3500P V5 specifications (LBA format: 4096 B)
11. Data retention period after power-off within the life cycle
12. Data retention period after power-off within the first 3% of the life cycle

4.1.2 ES3600P V5

Table 4-2 ES3600P V5 specifications (LBA format: 512 B)

Feature	Specifications			
Model	HWE52P43800 M002N	HWE52P431T6 M002N	HWE52P433T2 M002N	HWE52P436T4 M002N
Product name	ES3600P V5-800	ES3600P V5-1600	ES3600P V5-3200	ES3600P V5-6400

Feature	Specifications			
Form	2.5-inch disk PCIe 3.0	2.5-inch disk PCIe 3.0	2.5-inch disk PCIe 3.0	2.5-inch disk PCIe 3.0
Capacity	800 GB	1600 GB	3200 GB	6400 GB
Flash chip type	3D TLC	3D TLC	3D TLC	3D TLC
Sequential read bandwidth (MB/s)	3,500	3,500	3,500	3,500
4KB random read IOPS (Steady)	420,000	740,000	825,000	825,000
Average read latency@1QD (μs)	96	96	96	88
Sequential write bandwidth (MB/s)	1,000	1,900	3,200	3,050
4KB random write IOPS (steady/Max.)	115,000	195,000	300,000	250,000
Average write latency@1QD (μs)	15	14	14	15
Hybrid (R/W: 7/3) random 4KB IOPS (steady)	195,000	320,000	435,000	435,000
PBW (lifetime random PB written)	4.38 PB	8.76 PB	17.52 PB	35.04 PB
DWPD (JESD219)	3	3	3	3
Average power consumption (idle, activity)	5 W, 13 W	5 W, 13.8 W	5.5 W, 18.8 W	5.7 W, 21 W
Weight (g)	182			
Power-off protection	Supported			
Chip failure protection	Supported			
Trim	Supported			

Feature	Specifications
Mean time between failures (MTBF)	2.5 million hours
Annual failure rate (AFR)	≤ 0.35%
Uncorrectable bit error rate (UBER)	10^{-18}
Data Retention (power off)	3 months at 40 ℃ (104 ℉)
Data retention of a new SSD used a spare (power off)[12]	12 months at 40℃ (104 ℉)

NOTE
Detailed refers to ES3500P V5's specification notes.

4.1.3 ES3600C V5

Table 4-3 ES3600C V5 specifications (LBA format: 512 B)

Item	Specifications			
Model	HWE56P43800M002N	HWE56P431T6M002N	HWE56P433T2M002N	HWE56P436T4M002N
Product name	ES3600C V5-800	ES3600C V5-1600	ES3600C V5-3200	ES3600C V5-6400
Form	HHHL PCIe card, PCIe 3.0	HHHL PCIe card, PCIe 3.0	HHHL PCIe card, PCIe 3.0	HHHL PCIe card, PCIe 3.0
Capacity	800 GB	1600 GB	3200 GB	6400 GB
Flash chip type	3D TLC	3D TLC	3D TLC	3D TLC
Sequential read bandwidth (MB/s)	3,500	3,500	3,500	3,500
4 KB random read IOPS (steady)	420,000	740,000	825,000	825,000
Average read latency @1QD (μs)	96	96	96	88

Item	Specifications			
Sequential write bandwidth (MB/s)	1,000	1,900	3,200	3,050
4 KB random write IOPS (steady/maximum)	115,000	195,000	300,000	250,000
Average write latency @1QD (μs)	15	14	14	15
Hybrid (R/W: 7/3) random 4 KB IOPS (steady)	195,000	320,000	435,000	435,000
PBW (Lifetime Random PB Written)	4.38 PB	8.76 PB	17.52 PB	35.04 PB
DWPD (JESD219)	3	3	3	3
Average power consumption (Idle, Active)	5 W, 13 W	5 W, 13.8 W	5.5 W, 18.8 W	5.7 W, 21 W
Weight (g)	182			
Power failure protection	Supported			
Granular failure protection	Supported			
Trim	Supported			
Mean Time Between Failures (MTBF)	2.5 million hours			
Annual failure rate (AFR)	≤ 0.35%			
Uncorrectable bit error rate (UBER)	10^{-18}			
Data retention period (power failure)	3 months at 40℃			
Data retention period when	12 months at 40℃			

Item	Specifications
new disks are used as spare parts (power failure)	

NOTE
For details, see ES3500P V5 specifications notes.

4.2 Capacity

ES3500P V5	User Addressable Sectors in LBA Mode
1000 GB	1,953,525,168
2000 GB	3,907,029,168
4000 GB	7,814,037,168
8000 GB	15,628,053,168

ES3600P V5	User Addressable Sectors in LBA Mode
800 GB	1,562,824,368
1600 GB	3,125,627,568
3200 GB	6,251,233,968
6400 GB	12,502,446,768

ES3600C V5	User Addressable Sectors in LBA Mode
800 GB	1,562,824,368
1600 GB	3,125,627,568
3200 GB	6,251,233,968
6400 GB	12,502,446,768

NOTE
1 GB = 1,000,000,000 bytes; 1 sector = 512 bytes or 512 + 8 bytes.

4.3 Quality of Service (QoS)

Table 4-4 ES3500P V5 QoS (LBAformat: 512 B)

Specification	Unit	ES3500P V5			
		QD = 1			
		1000 GB	2000 GB	4000 GB	8000 GB
Quality of Service[1,2] (99.9%)					
Reads	µs	109	109	109	128
Writes	µs	16	14	14	14
Quality of Service[1,2] (99.99%)					
Reads	µs	166	182	223	197
Writes	µs	19	17	17	17

Table 4-5 ES3600P V5 QoS (LBA format: 512 B)

Specification[1,2]	Unit	ES3600P V5			
		QD = 1			
		800 GB	1600 GB	3200 GB	6400 GB
Quality of Service[1,2] (99.9%)					
Reads	µs	109	109	109	128
Writes	µs	14	14	14	16
Quality of Service[1,2] (99.99%)					
Reads	µs	142	157	206	239
Writes	µs	100	17	17	17

Table 4-6 ES3600C V5 QoS (LBA format: 512 B)

Specification[1,2]	Unit	ES3600C V5			
		QD = 1			
		800 GB	1600 GB	3200 GB	6400 GB
Quality of Service[1,2] (99.9%)					
Reads	µs	109	109	109	128
Writes	µs	14	14	14	16

Quality of Service[1,2] (99.99%)					
Reads	µs	142	157	206	239
Writes	µs	100	17	17	17

NOTE

1. Device measured using FIO. Quality of Service measured using 4 KB (4,096 bytes) transfer size on a random workload on a full Logical Block Address (LBA) span of the disk once the workload has reached steady state.
2. When the queue depth is 1, perform the 4 KB random I/O read/write test to collect statistics on the I/O command response time.

4.4 IOPS Consistency

Table 4-7 ES3500P V5 IOPS consistency (LBA format: 512 B)

Specification[1]	Unit	ES3500P V5			
		QD = 1			
		1000 GB	2000 GB	4000 GB	8000 GB
Random 4KB Read (up to)[2]	%	98	98	98	98
Random 4KB write (up to)	%	95	95	95	95

Table 4-8 ES3600P V5 IOPS consistency (LBA format: 512 B)

Specification[1]	Unit	ES3600P V5			
		QD = 1			
		800 GB	1600 GB	3200 GB	6400 GB
Random 4KB Read (up to)[2]	%	98	98	98	98
Random 4KB Write (up to)	%	92	95	95	92

Table 4-9 ES3600C V5 IOPS consistency (LBA format: 512 B)

Specification[1]	Unit	ES3600C V5			
		800 GB	1600 GB	3200 GB	6400 GB
Random 4KB Read	%	98	98	98	98

(up to)[2,3]					
Random 4KB Write (up to)[2,3]	%	92	95	95	92

NOTE

- The IOPS is sampled per second during the performance test. The sampling data is sorted in descending order. The minimum IOPS in the first 99.9% data is compared with the average IOPS during the test. Measurements are performed on a full Logical Block Address (LBA) span of the disk once the workload has reached steady state.
- OPS consistency is the result of testing 4 KB random I/Os by using the FIO tool in 8-thread 64QD mode.
- 4 KB = 4,096 bytes

4.5 Hot Swap

Hot swap includes the following two ways:

- Orderly hot swap: Users can directly insert a PCIe SSD when the OS is running, but needs to notify the OS before removing a PCIe SSD.
- Surprise hot swap: Users can directly insert or remove a PCIe SSD.

CAUTION

Before performing surprise hot swap operations, ensure that the following conditions are met:

- The ES3000 V5 is installed on a V5 server.
- The VMD function has been enabled in the BIOS.

ES3500P V5 and ES3600P V5 which are 2.5-inch form factor NVMe SSD supports orderly hot swap and surprise hot swap.

ES3600C V5 supports orderly hot swap.

4.6 NVMe Feature

Huawei ES3000 V5 SSD supports the NVMe 1.3 specification. Besides the basic feature and the mandatory commands, it also supports Name Space feature, and a maximum of 64 Name Spaces are supported.

4.7 Environmental Conditions

Table 4-10 2.5-inch SSD environmental specifications

Item	Description

Item	Description
Product	ES3500P V5 and ES3600P V5
Temperature	Storage temperature: –40 ℃ to +85 ℃ (–40 ℉ to +185 ℉) Operating temperature (Case): 0 ℃ to +70 ℃ (32 ℉ to +158 ℉)[1] Environmental temperature: 0 ℃ to +55 ℃ (32 ℉ to +131 ℉)
Altitude	• Operating: –305 m to 5486 m (–1,001 ft to 17998.50 ft) • Non-operating: –305 m to 12,192 m (–1,001 ft to 40,000 ft) The operating temperature decreases by 1 ℃ (1.8 ℉) with every increase of 220 m (721.78 ft) from 1800 m (5905.47 ft).
Humidity	• Operating: 5% to 95% • Non-operating: 5% to 95%
Shock	• Operating: 1000 G (Max) at 0.5 ms • Non-operating: 1000 G (Max) at 0.5 ms
Vibration	• Operating: 2.17 GRMS (5–700 Hz) Max • Non-operating: 3.13 GRMS (5–800 Hz) Max

NOTE

1. Operating temperature implies ambient air temperature under defined airflow in Table 4-12.

Table 4-11 SSD card environmental specifications

Item	Description
Product	ES3600C V5
Temperature	Storage temperature: –40 ℃ to +85 ℃ (–40 ℉ to +185 ℉) Operating temperature (Case): 0 ℃ to +70 ℃ (32 ℉ to +158 ℉)[1] Environmental temperature: 0 ℃ to +55 ℃ (32 ℉ to +131 ℉)
Altitude	• Operating: –305 m to 5486 m (–1,001 ft to 17998.50 ft) • Non-operating: –305 m to 12,192 m (–1,001 ft to 40,000 ft) The operating temperature decreases by 1 ℃ (1.8 ℉) with every increase of 220 m (721.78 ft) from 1800 m (5905.47 ft).
Humidity	• Operating: 5% to 95% • Non-operating: 5% to 95%
Shock	• Operating: 31 G at 2 ms, only one impact on the device's performance surface • Non-operating: 70 G at 2 ms, only one impact on the device's performance surface
Vibration	• Operating: 0.27 GRMS (5–500 Hz), 3 axial directions and 10 minutes for each axial direction • Non-operating: 1.49 GRMS (10–500 Hz), 3 axial directions and 15

Item	Description
	minutes for each axial direction

4.8 Thermal Specification

4.8.1 Boundary Conditions

The ES3000 V5 2.5-inch SSD is designed to work in an environment where air is provided at a typical temperature with specified air flow. The specified air flow and temperature is defined in Table 4-12.

The ES3000 V5 2.5-inch SSDs can work not only in servers that adopt front-to-rear air flow (type air flow for rack servers), but also in those that adopt horizontal air flow, as shown in Figure 4-1.

Figure 4-1 Design airflow direction

4.8.2 Thermal Specification

Table 4-12 2.5-inch SSD thermal specification

Name	TDP	Airflow	Operation condition
ES3600P V5-800	13 W	155LFM	Ambient temperature: 0 to 40 ℃ (32 ℉ to 104 ℉)

Name	TDP	Airflow	Operation condition
		650LFM	Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3600P V5-1600	13.8 W	155LFM 650LFM	Ambient temperature: 0 to 40 ℃ (32 ℉ to 104 ℉) Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3600P V5-3200	18.8 W	237LFM 650LFM	Ambient temperature: 0 to 40 ℃ (32 ℉ to 104 ℉) Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3600P V5-6400	21 W	237LFM 650LFM	Ambient temperature: 0 to 40 ℃ (32 ℉ to 104 ℉) Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3500P V5-1000	13 W	155LFM 650LFM	Ambient temperature: 0 to 40 ℃ (32 ℉ to 104 ℉) Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3500P V5-2000	13.8 W	155LFM 650LFM	Ambient temperature: 0 to 40 ℃ (32 ℉ to 104 ℉) Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3500P V5-4000	18.8 W	237LFM 650LFM	Ambient temperature: 0 to 40 ℃ (32 ℉ to 104 ℉) Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3500P V5-8000	21 W	237LFM 650LFM	Ambient temperature: 0 to 40 ℃ (32 ℉ to 104 ℉) Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)

Table 4-13 SSD card Sthermal specification

Name	TDP	Airflow	Operation condition
ES3600C V5-800	13 W	300LFM	Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3600C V5-1600	13.8 W	300LFM	Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3600C V5-3200	18.8 W	300LFM	Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)
ES3600C V5-6400	21 W	450LFM	Ambient temperature: 0 to 55 ℃ (32 ℉ to 131 ℉)

4.8.3 Detecting and Self-Protection

ES3000V5 will detect the SSD controller’s junction temperature and flash ambient temperature; SMART information will record the maximum temperature SSD has reached and the number and time of over temperature.

The maximum junction temperature of NAND flash used by ES3000 V5 can be 85 ℃ (185 ℉). According to the test result of Huawei Media Laboratory: below 78 ℃ (172.4 ℉), the temperature threshold of NAND Flash, SSD can work reliably in the whole lifecycle; over 78 ℃ (172.4 ℉), NAND flash can access and store the data correctly, but running for a long time in this environment[1], the lifecycle of the SSD will be affected; over 85 ℃ (185 ℉), the NAND flash cannot work properly. Therefore, Huawei designed the following temperature control strategy:

- In-band, firmware implements the alarms function and performance limit function by detecting the inner temperature.
- Out-of-band, by detecting the highest temperature of NAND flash in the SSD device, BMC adjust the wind speed to avoid triggering the temperature control mechanism.

For in-band management, firmware uses the 2 level alarm mechanism:

The first warning threshold is 78 ℃ (172.4 ℉). When the temperature reaches this threshold, the device will send alarm information to remind the user that the equipment overheating.

The second alarm threshold is 85 ℃ (185 ℉). When the temperature reaches the second threshold, the firmware will limit the performance of the device.

When the temperature reaches the second threshold and the overtemperature duration is more than 6 minutes, to ensure data reliability, ES3000 V5 will further restrict read and write performance, lower power consumption and to avoid exceeding the specifications of the products. When the temperature is reduced below the RT level, the performance limit will be lifted.

Table 4-14 Temperature self-protection design

Grade	Threshold	Action
OT Level 1[2]	composite: 78 ℃ (172.4 ℉)	Warning.
RT Level 1[3]	composite: 73 ℃ (163.4 ℉)	The warning will be lifted.
OT Level 2[2]	composite: 85 ℃ (185 ℉)	Warning, limit the performance.
RT Level 2[3]	composite: 78 ℃ (172.4 ℉)	The performance limit will be lifted, and the warning remains active.

NOTE

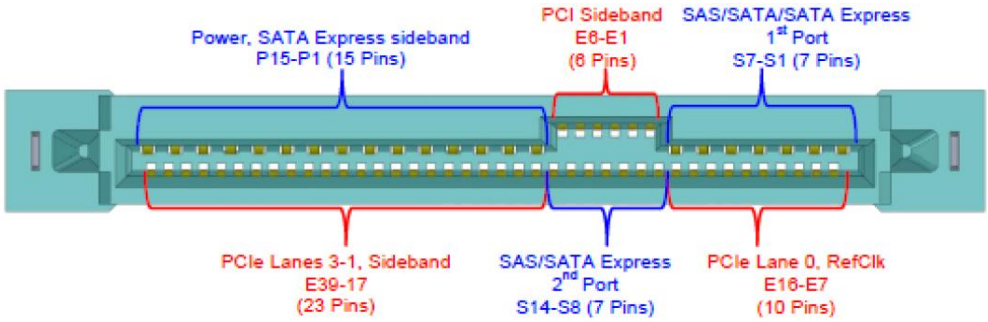
1. The SSD can work at 85 ℃ (185 ℉) for a maximum of 6 minutes.
2. OT: Over Temperature
3. RT: Recovery temperature

5 Pin and Signal Descriptions

5.1 Pin and Signal for 2.5-inch SSD

The connector of 2.5-inch SSD is SFF-8639. The PCIe signal assignments are as follows:

Figure 5-1 Signal Assignments



The pin definitions are as follows:

Table 5-1 Pin definition (Compatible with SFF-8639)

Pin	Name	Description	Pin	Name	Description
S1	GND	Ground	E7	RefClk0+	ePCIe Primary RefClk +
S2		NC	E8	RefClk0-	ePCIe Primary RefClk-
S3		NC	E9	GND	Ground

Pin	Name	Description	Pin	Name	Description
S4	GND	Ground	E10	PETp0	ePCIe 0 Transmit +
S5		NC	E11	PETn0	ePCIe 0 Transmit -
S6		NC	E12	GND	Ground
S7	GND	Ground	E13	PERn0	ePCIe 0 Receive -
E1	REFCLK1+	Reference clock port B (not used)	E14	PERp0	ePCIe 0 Receive +
E2	REFCLK1-	Reference clock port B (not used)	E15	GND	Ground
E3	3.3Vaux	3.3 V for SM bus	E16	RSVD	Reserved
E4	ePERst1#	ePCIe Rest (port B)(not used)	S8	GND	Ground
E5	ePERst0#	ePCIe Rest (port A)	S9		NC
E6	RSVD (Vendor)	Reserved (vendor)	S10		NC
P1		NC	S11	GND	Ground
P2		NC	S12		NC
P3		NC	S13		NC
P4	IfDet#	Interface Detect (Was GND-percharge)	S14	GND	Ground
P5	GND	Ground	S15	RSVD	Reserved
P6	GND	Ground	S16	GND	Ground
P7		NC	S17	PETp1	ePCIe 1 Transmit +
P8			S18	PETn1	ePCIe 1 Transmit -
P9			S19	GND	Ground

Pin	Name	Description	Pin	Name	Description
P10	PRSNT#	Presence (Drive type)	S20	PERn1	ePCIe1 Receive-
P11	Activity	Activity (output)/Spinup	S21	PERp1	ePCIe1 Receive+
P12	GND	Hot swap Ground	S22	GND	Ground
P13	12 V	Precharge All -12 V Only power for ePCIe SSD	S23	PETp2	ePCIe 2 Transmit+
P14			S24	PETn2	ePCIe 2 Transmit-
P15			S25	GND	Ground
			S26	PERn2	ePCIe2 Receive-
			S27	PERn2	ePCIe2 Receive+
			S28	GND	Ground
			E17	PETp3	ePCIe 3 Transmit +
			E18	PETn3	ePCIe 3 Transmit -
			E19	GND	Ground
			E20	PERn3	ePCIe 3 Receive-
			E21	PERp3	ePCIe 3 Receive+
			E22	GND	Ground
			E23	SMClk	SM-Bus Clock
			E24	SMDat	SM-Bus Data
			E25	DualPortEn#	ePCIe 2x2 Select

5.2 Pin and Signal for AIC SSD

	Side B		**Side A**	
Pin	Name	Description	Name	Description
1	+12 V	12 V power	PRSNT1#	Hot swap presence detection
2	+12 V	12 V power	+12 V	12 V power
3	+12 V	12 V power	+12 V	12 V power
4	GND	Ground	GND	Ground
5	SMCLK	SMBus clock	JTAG2	TCK (Test Clock), clock input for JTAG interface
6	SMDAT	SMbus data	JTAG3	TDI (Test Data Input)
7	GND	Ground/UART_HOST	JTAG4	TDO (Test Data Output)
8	+3.3 V	3.3 V power	JTAG5	TMS (Test Mode Select)
9	JTAG1	TRST# (Test Reset) resets the JTAG interface	+3.3 V	3.3 V power
10	3.3Vanx	3.3 V auxiliary power	+3.3 V	3.3 V 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 swap	PERn0	Receiver

	Side B		**Side A**	
		presence detection		differential pair, Lane 0
18	GND	Ground	GND	Ground
End of the x1 Connector				
19	PETp1	Transmitter differential pair, Lane 1	RSVD	
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
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 swap presence detection	GND	Ground

	Side B		**Side A**	
32	GND	Ground	RSVD	Reserved
End of the x4 Connector				

6 Management

ES3000 V5 provides maintenance and management functions with comprehensive functions and high performance. That includes the in-band management tools and the out-of-band management interfaces.

6.1 In-band Management

- Provide whole-disk wiping to ensure data confidentiality when the SSD is sent for repair.
- Supports in-band online upgrades to facilitate routine maintenance.
- Provides device information in a centralized manner, including the model, capacities, temperature, remaining service life, and health status
- Supports log query for analyzing device health status.
- Provides the manufacture dates and serial numbers to facilitate asset management.
- Provides statistics functions to allow maintenance engineers to query the wear status, the ratio of bad blocks.

Besides, ES3000 V5 provides the GUI management tools in Windows platform, which facilitates the user's daily operation.

6.2 Out-of-band Management

Based on the SMBUS interface, the ES3000 V5 implements the out-of-band management function of *NVM Express Basic Management Command* in *NVM Express Management Interface Revision 1.0*. By means of SMBUS, BMC can access the VPD information in the EEPROM and read the management data of the NVMe SSD. VPD information of ES3000 V5 complies with the regulation *NVM Express Management Interface Revision 1.0*.

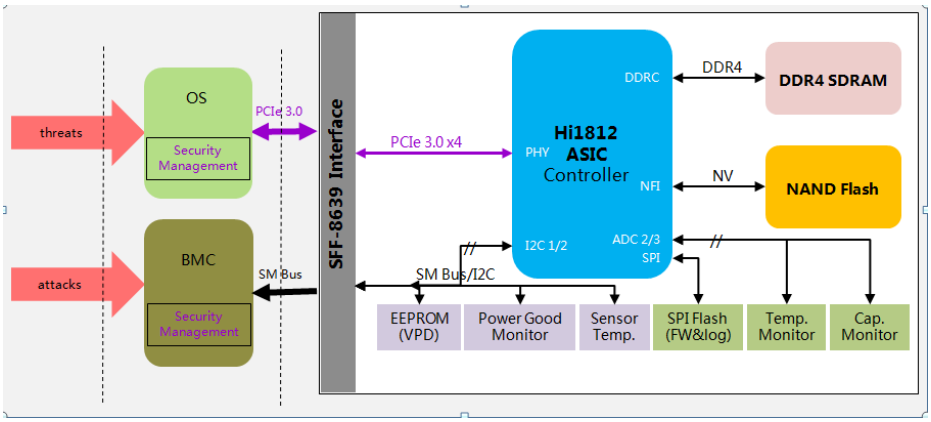
By means of out-of-band management interface, BMC can obtain the information including model number, healthy status, remaining life, firmware version, and so on.

Out-of-band management information can be accessed through address 0x6A (not include R/W bit). VPD's IIC address is 0x53 (not include R/W bit).

7 Security

7.1 Security Architecture

The following figure shows the security architecture based on the logical position of the ES3000 V5 in real applications. As shown in the figure, the ES3000 V5 SAS ports comply with NVMe 1.3 specifications. Various OSs perform data read/write through standard drive I/O ports, and the iBMC implements out-of-band management according to *NVM Express Management Interface Revision 1.0*. The standard drive I/O ports and out-of-band management rely on the security mechanisms of the OS and iBMC. Therefore, The ES3000 V5 has low security risks.



7.2 Security Risks and Threats

The following risk assessment process is used to identify security threats to the ES3000 V5.

The following table describes the security risk assessment for the related product components.

Main Component	Description	Risks Involved	Severity
NVMe driver	SSD kernel driver for interaction with the block file system and communication with the user space.	Yes	Low
CLI	Enables users to run commands to obtain SSD information, such as health status and temperature.	Yes	Low
SNMP agent	Enables users to manage SSDs over the network in a centralized manner, for example, obtaining the SSD health status and temperature.	Yes	Low
Firmware	As the front-end protocol processing module of the NVMe device, the firmware processes I/O between the device and the host by transmitting and parsing data.	None	
BMC interface	Enables users to manage SSDs through the BMC, for example, obtaining the SSD health status and temperature.	Yes	Low

For details about risk solutions, see the next section.

7.3 Security Policies and Solutions

Huawei is committed to safeguarding the stability and security of customers' networks and services.

As a storage device in a server or storage host, the ES3000 V5 follows the lowest permission security policies and provides the following functions to enhance security:

- Access control: Administrator permissions are required to use any device maintenance and management tools of the ES3000 V5.
- Operation log: All non-query operations performed on the ES3000 V5 are logged.
- Security assurance: A complete data clearance tool is provided to destroy data at the NAND flash memory chip level, preventing data leakage.
- Firmware upgrade: Firmware upgrade through in-band management channels ensures upgrade reliability and network security. The upgrade tool performs an upgrade only after the upgrade file passes the integrity and validity checks. Firmware can be loaded only after it is authenticated by the Hi1812E controller, which ensures firmware validity and data security.

- Device management: The ES3000 V5 uses out-of-band management based on I2C channels and in-band management based on SNMP. Both management functions support only query, preventing malicious and illegal write operations.

8 Maintenance

Visit

http://support.huawei.com/carrier/docview!docview?nid=DOC1100361702.

9 Certifications

Table 9-1 lists the certifications that the ES3000 V5 has passed.

Table 9-1 Certifications

No.	Country/Region	Certification
1	China	RoHS
2	Europe	WEEE
3	Europe	RoHS
4	Europe	REACH
5	Europe	CE
6	America	FCC
7	America	NRTL
8	America	IC
9	Japan	VCCI
10	Australia	RCM
11	Saudi Arabia	KSA
12	Kuwait	KUCAS
13	Nigeria	SONCAP
14	Uganda	UNBS
15	Alegria	ACAP
16	Member States of IEEE	CB

10 DVT and Qualification

10.1 Signal Integrity

10.1.1 PCIe

Table 10-1 High speed PCIe link test

Item	Description	Result
TX Conformance	With SigTest tools, in accordance with the requirements of the PCIE specification, test PCIe TX signal including Jitter, Eye and other indicators Test controller chip samples cover Fast, Slow, and Typical chips	Pass
RX Tolerance	With BertScope tools, in accordance with the requirements of the PCIE specification, test PCIe RX tolerance Test controller chip samples cover Fast, Slow, and Typical chips	Pass
Pull off	Measuring the quality of PCIe signals for the longest and shortest link Test controller chip samples cover Fast, Slow, and Typical chips	Pass
Build chain	Repeatedly to build chain. Test covers the longest and shortest link Test chip samples cover Fast, Slow, and Typical chips	Pass
PVT (Process, Voltage,	Operating temperature (Case) for 2.5-inch: 0 to 70 ℃	Pass

Item	Description	Result
Temperature)	(32 ℉ to 158 ℉) Operating temperature (Ambient) for AIC: 0 to 55 ℃ (32 ℉ to 131 ℉) Voltage deviation: +/–5% Controller Corner ship: FF, SS, and TT.	

10.1.2 Input Clock

In accordance with the PCIE specification, inject jitter and frequency offset with the clock generator, test the input clock tolerance.

Test samples cover PVT.

10.2 Protocol Compliance

10.2.1 PCIe

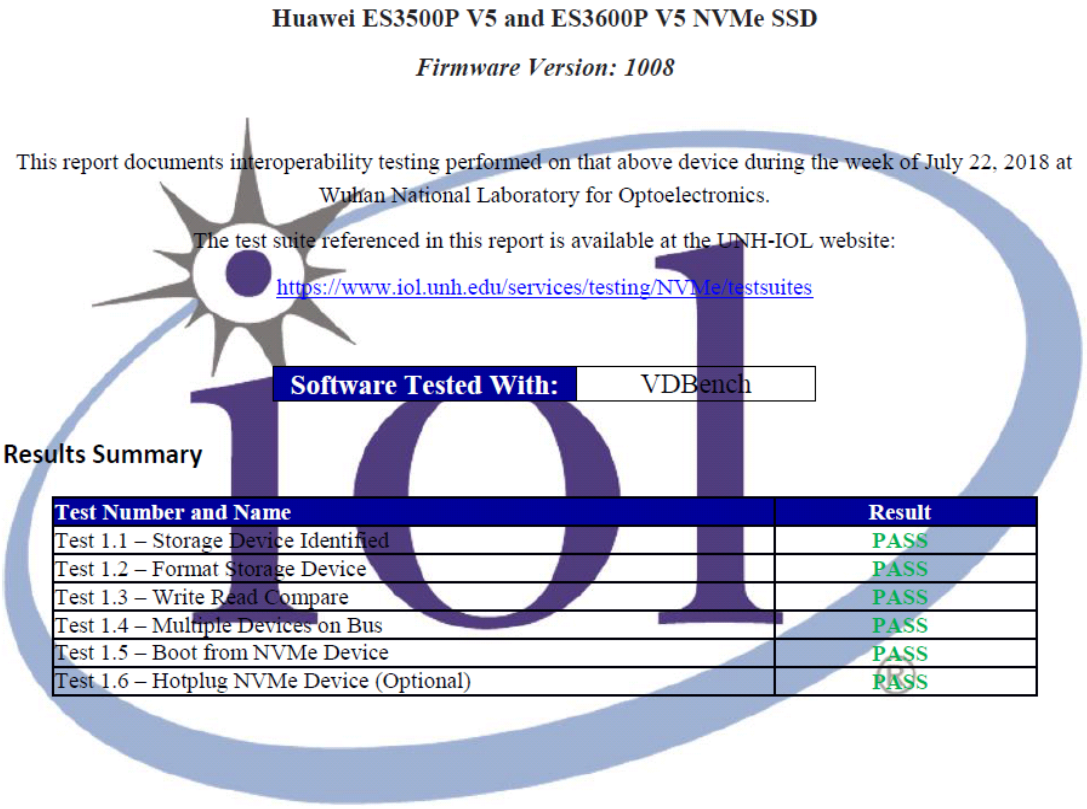
ES3000 V5 will participate in the PCI-SIG certification test in July 2018.

10.2.2 NVMe

ES3000 V5 supports the NVMe 1.3 protocol and has passed the NVMe Plufest test on July 22, 2018. The test result shows that NVMe SSDs of the ES3500P V5 and ES3600P V5 series have been added to the NVMe interworking product list.

The following figure shows the test result (NVMe 1.3):

Figure 10-1 Test result (NVMe 1.3)



Huawei ES3500P V5 and ES3600P V5 NVMe SSD

Firmware Version: 1008

This report documents interoperability testing performed on that above device during the week of July 22, 2018 at Wuhan National Laboratory for Optoelectronics.

The test suite referenced in this report is available at the UNH-IOL website:

https://www.iol.unh.edu/services/testing/NVMe/testsuites

Software Tested With:	VDBench

Results Summary

Test Number and Name	Result
Test 1.1 – Storage Device Identified	PASS
Test 1.2 – Format Storage Device	PASS
Test 1.3 – Write Read Compare	PASS
Test 1.4 – Multiple Devices on Bus	PASS
Test 1.5 – Boot from NVMe Device	PASS
Test 1.6 – Hotplug NVMe Device (Optional)	PASS

10.2.3 SNMP

Net-SNMP is open source code software that supports SNMPv1, SNMPv2c, SNMPv3. SNMPv3 is recommended due to its high security. Net-SNMP includes the source codes of the agent and multiple management tools and supports multiple extension modes.

The SNMP service utility of Windows provides SNMP query and trap functions, but supports only SNMPv1 and SNMPv2c. SNMPv3 is not supported. Therefore, exercise caution when using the utility.

The ES3000 V5 provides the SNMP extension agent to support the SNMP function. The extension defines the device Management Information Base (MIB). SNMP clients can query device status through the SNMP extension agent.

10.3 Functional and Interoperability

Table 10-2 Functional and interoperability test list

No.	Item	Description	Result
1	Installation	OS: Linux/Windows/VMware 20+ times test for each compatible OS	Pass
2	Firmware	4 samples for each specification	Pass

No.	Item	Description	Result
	Update	Upgrade/rollback 20+ times for each test device	
3	Bootable	4 samples for each specification Data conformance testing Reboot and power on/off test	Pass
4	Performance	4 samples for each specification Bandwidth, IOPS, latency, stability, QoS	Pass
5	Power Cycle	4 samples for each specification 2000+ times for each test device	Pass
6	Reboot	4 samples for each specification 2000+ times for each test device	Pass
7	Power loss	4 samples for each specification 2000+ times for each test device	Pass
8	Hot swap	4 samples for each specification 300+ times for each test device in capable platforms and OSs	Pass
9	Fault Injection	Injection coverage: disk, Flash, and PCIe	Pass

10.4 Reliability

Reliability list

Item	Description	Result
RDT	Endurance test based on JESD218A/219 Standard 762 disks for RDT verification	Pass
HALT	Change the following conditions to accelerate the equipment aging. During the test, observe the functions and performance of the device. Server inlet temperature: –20 ℃ to +60 ℃ (–4 ℉ to +140 ℉) Voltage deviation: +/–5% ES3000 controller chip: FF, SS, and TT	Pass
HASS	The following conditions are changed to accelerate the exposure of potential defects of the device. During the test, observe the functions and performance of the device. Operating temperature (Case) for 2.5-inch: 0 to 65 ℃ (32 ℉ to 149 ℉)	Pass

Item	Description	Result
	Tests: Performance and power on/off test	
Long Term Stability	Sample: 60 disks Time: 3 months per sample Tests: data conformance testing	Pass