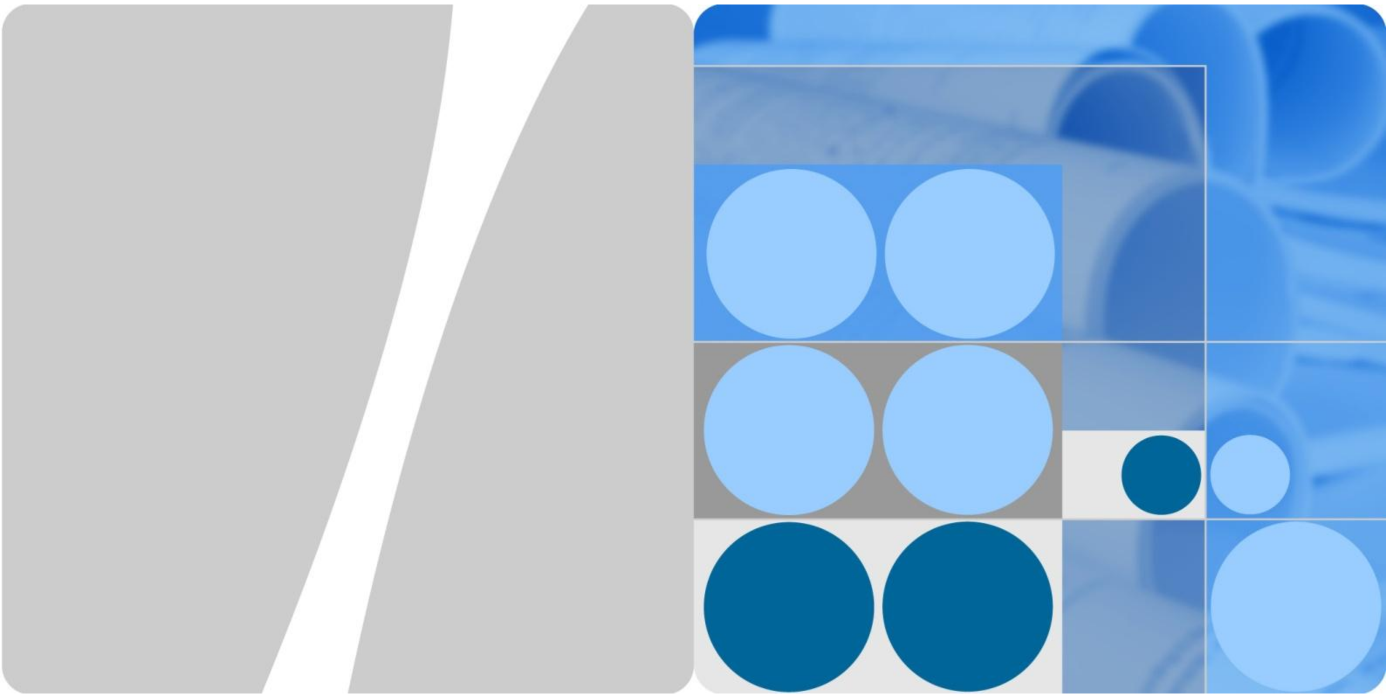


Huawei Enterprise SSD

Technical White Paper

Issue	02
Date	2017-04-20

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions

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://e.huawei.com

Contents

1 Overview

Storage technologies have gained fast development in the past two decades. Representing the computing capability, CPU performance has increased by nearly 580 times. I/O channel performance is also almost 1000 times higher than before. However, improvement in storage system media is only 20 times.

Hard disk drives (HDDs) have become a performance bottleneck for computer systems, greatly impeding the improvement of overall IT system performance and making the system unable to meet business development requirements.

Not having mechanical parts such as magnetic heads, spindles, rotating motors, and other moving parts that are integral to HDDs, solid state drives (SSDs) are free from mechanical faults and can work properly in the existence of collisions, shocks, and vibrations. Compared to HDDs, SSDs have absolute advantages in terms of performance, reliability, power consumption, and portability. These advantages have enabled SSDs to be widely adopted in various industries.

Huawei entered the solid state storage market in 2005 and has been focusing on enterprise storage applications. With accumulated expertise, Huawei has more than 150 patents in SSDs. Huawei's proprietary SSD product, Huawei Enterprise Solid State Drive , is designed for enterprise storage and has generation-over-generation optimizations. This document describes the basic principles, algorithms, and customer benefits of SSDs to help readers understand and sales personnel sell Enterprise Solid State Drive.

This document is intended for Huawei sales, service, R&D, and marketing personnel and Huawei customers.

2 Basic Working Principles

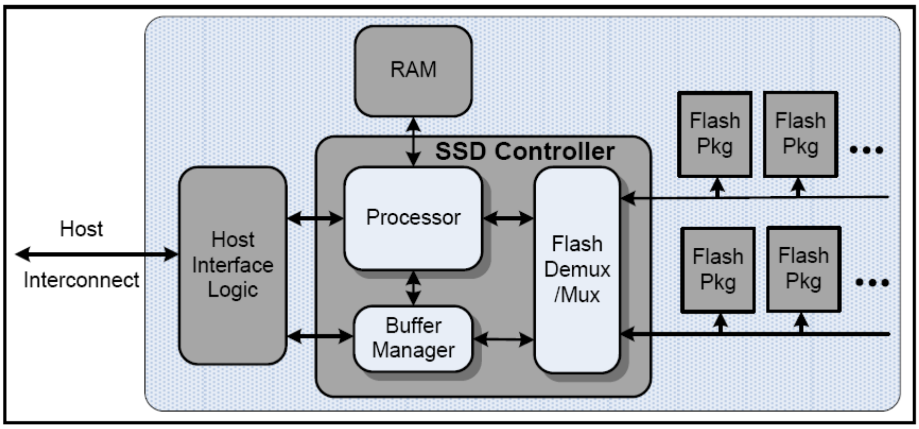
Unlike HDDs that adopt magnetic heads and disks, SSDs use electronic components such as SSD controllers and nonvolatile storage chips to store data. Despite this, SSDs are the same as HDDs in interface specifications, sizes, and ways of use. SSDs described in this document all use NAND flash as storage media.

2.1 SSD Architecture

SSDs consist of control units and storage units (mainly flash memory chips currently). Control units contain SSD controllers, host interfaces, and dynamic random access memory (DRAM) modules. Storage units contain only NAND flash chips.

- Host interface: the protocol and physical interface used by a host to access an SSD. Common interfaces are SATA, SAS, and PCIe.
- SSD controller: a core SSD component responsible for read and write access from a host to the back-end media and for protocol conversion, table entry management, data caching, and data checking.
- DRAM: a component responsible for the Flash Translation Layer (FTL) table and data caching to provide fast data access.
- NAND flash: a physical component for storing data. (For details, see the following sections.)

Figure 2-1 SSD architecture



Huawei Enterprise SSD use Huawei proprietary SSD controller chips.

2.2 NAND Flash

NAND flash is a nonvolatile random access storage media. Based on a floating gate transistor design, it stores electric charges by floating gates so that data can be stored even without an external power supply. In comparison to HDDs' storage media, NAND flash is rapid in data reading and writing and more responsive in access latency.

2.2.1 Working Principles

Data Storage

NAND flash stores data using floating gate transistors. The threshold voltage changes based on the number of electric charges stored in a floating gate. Data is then represented using the transistor's threshold voltage that is read.

Internal Structure of NAND Flash

Internal storage units in NAND flash include LUNs, planes, blocks, pages, and cells.

- LUN (or DIE): the minimum physical unit that can be independently encapsulated. A LUN typically contains multiple planes.
- Plane: a storage unit that has independent page registers. A plane typically contains 1k or 2k odd or even number of blocks.
- Block: the minimum unit that can be erased. A block typically contains multiple pages.
- Page: the minimum unit that can be programmed and read. Typical page sizes include 4 KB, 8 KB, and 16 KB.
- Cell: the minimum erasure, writing, and reading unit in a page. A cell corresponds to a floating gate transistor and can store one or more bits.

Figure 2-2 NAND flash structure

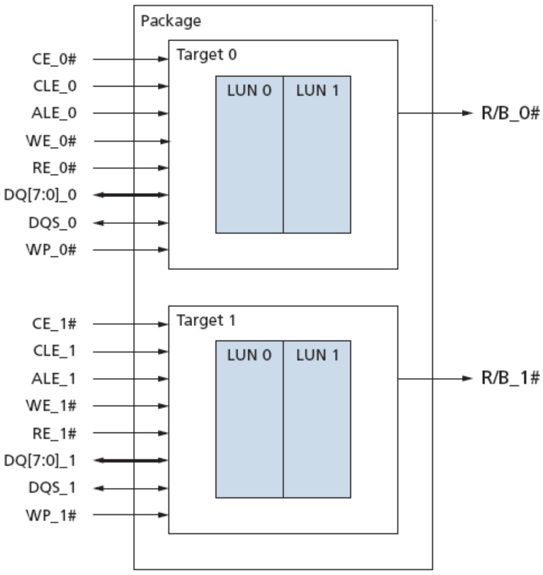
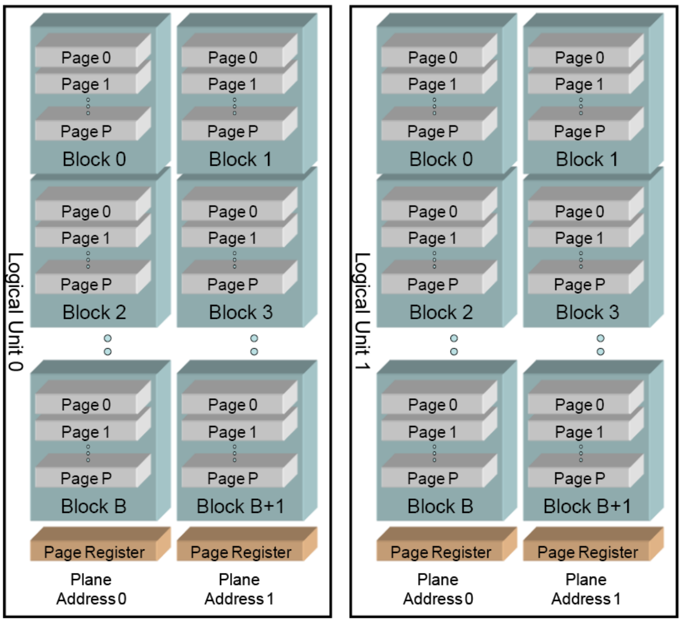
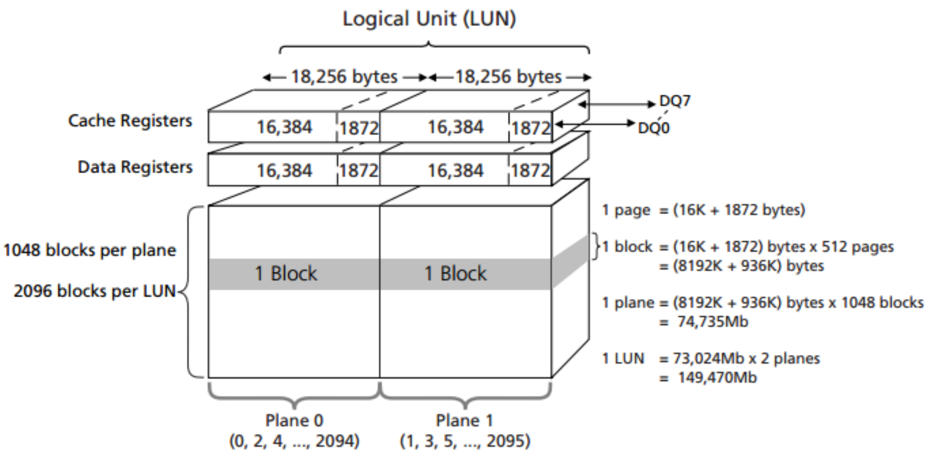


Figure 2-3 LUN structure



Data read and write operations on NAND flash mainly include erasure, programming, and reading. The basic unit for programming and reading is page and that for erasure is block. Because error checking and correction (ECC) need to be performed on data stored in NAND flash, page sizes are bigger than integral 4 KB, 8 KB, or 16 KB. For example, a 16-KB page's actual size may be 16,384+1872 bytes, where the 16,384 bytes are used for storing data and the 1872 bytes for storing ECC codes. The area used for ECC codes is called an Out-of-Bank (OOB) area.

Figure 2-4 NAND flash storage matrix structure (micron 16 nm MLC)


Logical Unit (LUN)
18,256 bytes
18,256 bytes
DQ7
DQ0
Cache Registers
16,384
1872
16,384
1872
Data Registers
16,384
1872
16,384
1872
1048 blocks per plane
2096 blocks per LUN
1 Block
1 Block
1 page = (16K + 1872 bytes)
1 block = (16K + 1872) bytes x 512 pages = (8192K + 936K) bytes
1 plane = (8192K + 936K) bytes x 1048 blocks = 74,735Mb
1 LUN = 73,024Mb x 2 planes = 149,470Mb
Plane 0 (0, 2, 4, ..., 2094)
Plane 1 (1, 3, 5, ..., 2095)

2.2.2 Key Parameters

Endurance

Endurance of NAND flash is a key reliability indicator to measure the Program/Erase (P/E) life of NAND flash. It is represented by the P/E cycle. NAND flash is nonvolatile media. Before new data is written to it, the target block must be erased; otherwise, error bits may occur. Each time a block is erased and then written, a P/E cycle is calculated.

During a P/E process, the insulation layer of the floating gate is damaged to a certain extent. If a P/E attempt fails, the corresponding block is tagged as a bad block. NAND flash manufacturers accept that the service life of NAND flash comes to an end if 3% of blocks are bad blocks.

Besides endurance, the service life of an SSD is subject also to ECC algorithms, over-provisioned space, and internal management algorithms.

BER

Due to its inherent characteristics, NAND flash has bit inverting at a certain probability, and this probability is measured by the bit error rate (BER). BER is divided into raw bit error rate (RBER) and uncorrectable bit error rate (UBER).

- RBER: It is the bit error rate before error checking and correction and reflects the primal reliability status of NAND flash. The higher the RBER, the worse the NAND flash reliability.
- UBER: It indicates the uncorrectable bit inversion errors on a codeword because those errors are beyond the capability of ECC algorithm. UBER is a main indicator that can tell the probability at which ECC errors occur on a codeword.

At present, most manufacturers use UBER to prove their products in ECC capability, which typically ranges from 10^{-16} to 10^{-13}.

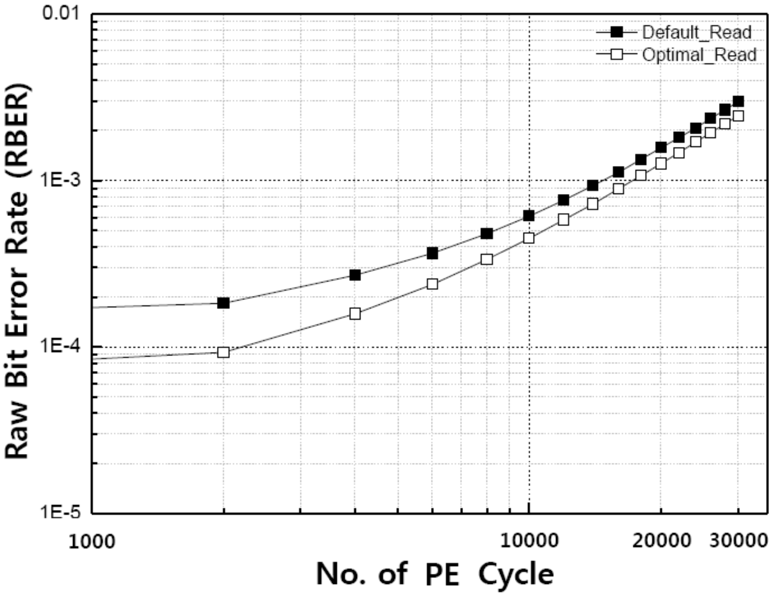
$$UBER = \frac{P_{cw}}{Data}$$

where $P_{cw} = \sum_{n=E+1}^{N} C_N^n RBER^n (1 - RBER)^{N-n}$ specifies the probability at which bit inverting occurs in a codeword at a certain point in time. Among the parameters, *n* specifies the number of bit inversions, *E* the maximum of bit inversions that can be corrected by ECC algorithm, *N* the total number of bits in a codeword, *RBER* the probability of a bit's inverting, and *Data* the number of user data bits in a codeword.

RBER increases as the P/E cycle count of NAND flash increases, for the oxide layer of the floating gate is worn during P/E operations. RBER also worsens as the process of NAND flash decreases. For example, there is a probability of 10^{-9} to 10^{-7} that a multi-level cell (MLC) of 50 nm process has RBER errors. For an MLC of 50 nm process, a general ECC algorithm is sufficient for ensuring data storage reliability. In comparison, for an MLC of 20 nm process with the probability of 10^{-5} to 10^{-3}, a more advanced ECC algorithm is required.

The RBER of flash memory chips is generally 10e-5, and becomes worse as the P/E count increases. To resolve this problem, the error correction algorithm of SSD controller chips must cover the whole SSD life cycle. Huawei uses the industry-leading low-density parity-check (LDPC) algorithm. LDPC is more efficient than Bose-Chaudhuri-Hocquenghem (BCH) under equivalent conditions. It helps minimize risks caused by the compromised reliability of flash memory chips.

Figure 2-5 P/E cycle and BER of a 20 nm MLC



Data Retention

Data retention of NAND flash is another major reliability indicator, which represents how long data can be stored in intact state in NAND flash. It means how long data can be integrally stored on NAND flash with ECC under a certain temperature range. There are two factors involved, P/E cycle count and ambient temperature. The data retention is inversely proportional to the P/E cycles and the ambient temperature. The higher the P/E cycles or the ambient temperature, the shorter the data retention time.

The data retention of enterprise SSDs is subject to the JEDEC JESD218 standard: After an enterprise SSD experiences a full P/E cycle and powers off in the room temperature of 40 ℃, the data retention time must be at least three months.

ECC

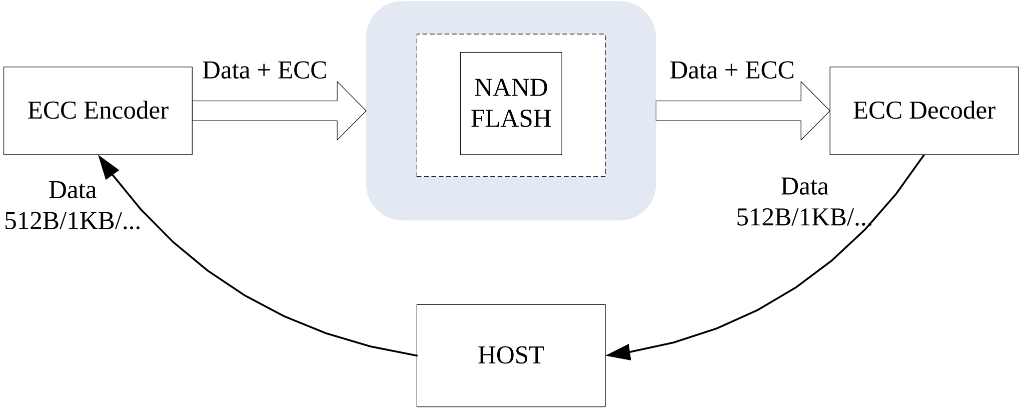
ECC stands for Error Correction Code. It is used to detect and correct errors. The types of ECC include Reed-Solomon code, Hamming code, BCH code, and LDPC code. Each type has its different application scenarios. LDPC code handles dispersed errors. It has a simple algorithm but powerful error correcting capabilities, and is easy to configure. LDPC code is widely applied in SSDs.

Common data check algorithms, such as parity check and cyclic redundancy check (CRC), adds extra data to the raw information data. Similarly, ECC also needs extra data space to save the check data generated by it. The extra OOB data on each page of NAND flash is the check code for ECC data. A larger OOB space means a more powerful ECC capability.

LDPC works on a per 4K-byte basis, which is also called a codeword. The 4K bytes are raw data, and LDPC must generate certain bits of check data. For example LDPC code whose bit rate is 90.00% for 4K bytes needs 455 bytes to store check code.

The working principle of ECC is as follows: Before the SSD controller writes data to NAND flash, the data is processed by the LDPC encoder. Then together with the check data, the raw data is written to NAND flash. When the SSD controller reads data from NAND flash, the raw data and the check data are directed to the LDPC decoder. Then the LDPC decoder uses the fixed-point Normalized Min-Sum Algorithm (NMSA) algorithm to correct errors. If the decoding succeeds, correct data is returned; if the decoding fails, other means will be used to read data.

Figure 2-6 Working principle of ECC



3 Core Technologies

3.1 SSD Controller

The SSD controller is the core component of an SSD. It determines the performance and reliability specifications of an SSD. Currently Huawei Enterprise SSD use the Huawei proprietary second-generation controller, which is designed for enterprise-class applications. This controller offers the industry's standard PCIe3.0 x4 and SAS3.0 x2 interfaces. It is advanced with its high performance and low power consumption as well as value-added storage features. To maximize the service life against media wearing, the SSD controller uses enhanced ECC and digital signal processing as well as built-in RAID to prolong the service life of SSDs, meeting the reliability requirements of enterprise-class applications. The SSD controller adopts the 28 nm process and supports DDR4 and SAS 12 Gbit/s, PCIe 8GT/s as well as hardware accelerated FTL, delivering stable performance with a short latency.

The SSD controller on Huawei Enterprise SSD has the following key features:

- Dual 12 Gbit/s SAS and PCIe 8GT/s ports
- 2D SLC/MLC/TLC or 3D MLC flash
- DDR4 interface
- 18 channels and 288 DIEs
- LDPC (hard/soft decoding)
- Cross-chip dynamic RAID
- Data optimization
- All-path data protection

3.2 Basic Algorithms

3.2.1 FTL Management

The flash translation layer (FTL) is a software layer allowing an SSD to simulate HDD operations. On HDDs, a piece of data can be overwritten on its physical location. Each piece of data logical block address (LBA) has a fixed physical block address (PBA). In comparison, data on a block of SSD NAND flash must be erased before new data is written to that block. Besides, data reads and writes on NAND flash are performed on a granularity of page and data erasure on a granularity of block (consisting of pages), leading to unfixed mapping between LBAs and PBAs. The FTL is used to manage the unfixed mapping. The FTL manages host data and arrange them with a certain order on each NAND flash chip of an SSD, maintains the mapping between LBAs and PBAs, and manages the status of blocks.

The FTL table is temporarily saved on the DDR, and periodically flushes updates to the NAND flash. When receiving a read instruction from a host, an HUAWEI ENTERPRISE SSD queries the FTL to obtain the data location on the NAND flash, and returns the requested data to the host. When receiving a write instruction, an HUAWEI ENTERPRISE SSD applies for a blank page and writes data to the page. Then it updates the FTL and redirects the LBA to that page.

3.2.2 Wear Leveling

The P/E cycle count for each NAND flash block is limited. Unchecked P/Es for a block will cause the block to malfunction earlier than other blocks. Too many malfunctioned blocks will cause the entire NAND flash to break down. The wear leveling algorithm is used to prevent unbalanced R/Es on a block. Wear leveling helps balance P/Es on NAND flash blocks, greatly lengthening the service life of NAND flash.

Wear leveling is classified into dynamic and static wear leveling. Dynamic wear leveling is triggered by host data changes and ensures that P/E requests are directed first to the blocks that are least worn. It balances P/E operations among each block. Static wear leveling periodically searches for the blocks that have the fewest P/E loads and reclaims data from those blocks. It ensures that the blocks that store cold data participate in wear leveling. Huawei Enterprise SSD combine dynamic and static wear leveling to balance loads on a whole HUAWEI ENTERPRISE SSD.

Figure 3-1 Working principle of dynamic wear leveling

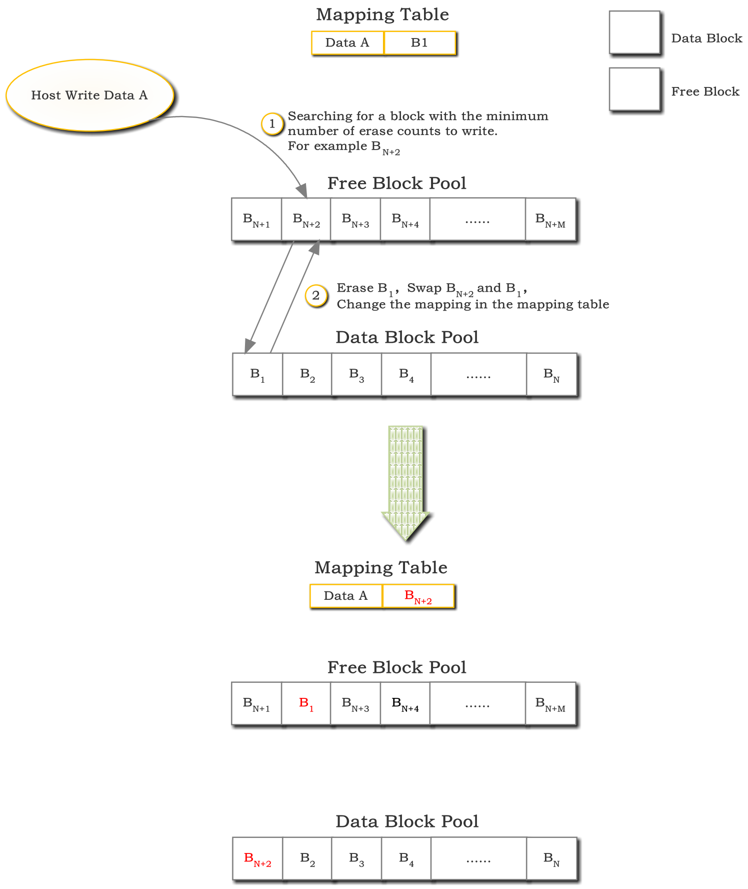
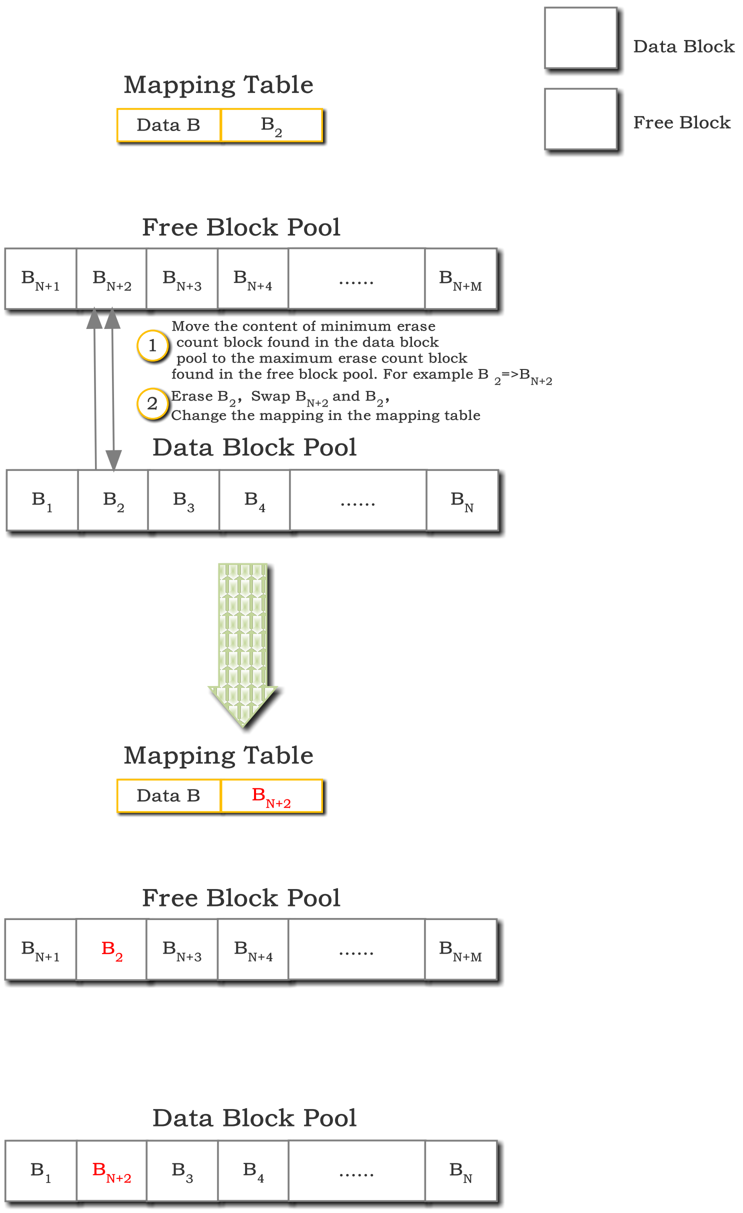


Figure 3-2 Working principle of static wear leveling

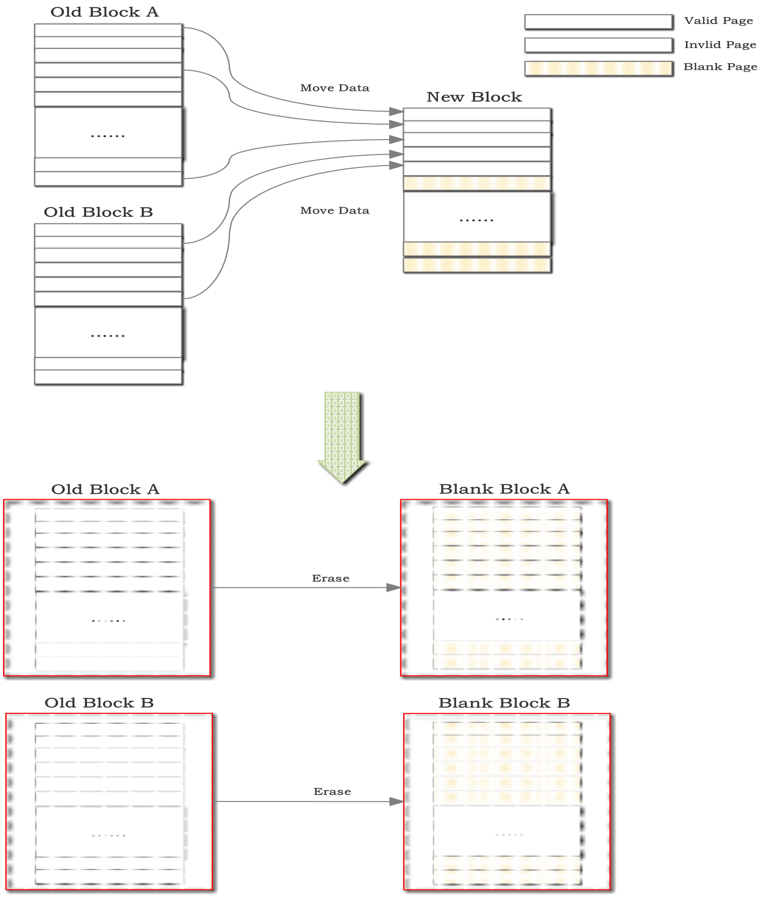


3.2.3 Garbage Collection

HDDs use magnetic media as storage media. Data on the magnetic media can be directly overwritten. SSDs use NAND flash as storage media. One significant feature of NAND flash is that data on the NAND flash must be erased before new data is written. A page is the basic write unit of NAND flash, and a block is a basic erase unit of NAND flash. A block consists of a certain number of pages. Before writing to a page, the system must erase the data on the block where the page is located. Data on SSDs is constantly updated and the NAND flash space is fixed. To ensure that there is always space to write data, old blocks must be released. Garbage collection helps serve this purpose.

When data is written to a page of a new block on an SSD, the page on the old block will become ineffective. If there are sufficient ineffective pages on the old block, you can reclaim the block to release space. The working principle of garbage collection is to move effective pages from old blocks to a new block and erase the old blocks to release space.

Figure 3-3 Working principle of garbage collection



Simply put, garbage collection is a process of converting ineffective blocks into empty blocks. Huawei Enterprise SSD adopt an advanced garbage collection algorithm that automatically optimizes garbage collection parameters based on the data write amount, service model and pressure, and the number of bad blocks. This algorithm minimizes ineffective data migrations and maximizes the SSD performance and service life.

3.2.4 Bad Block Management

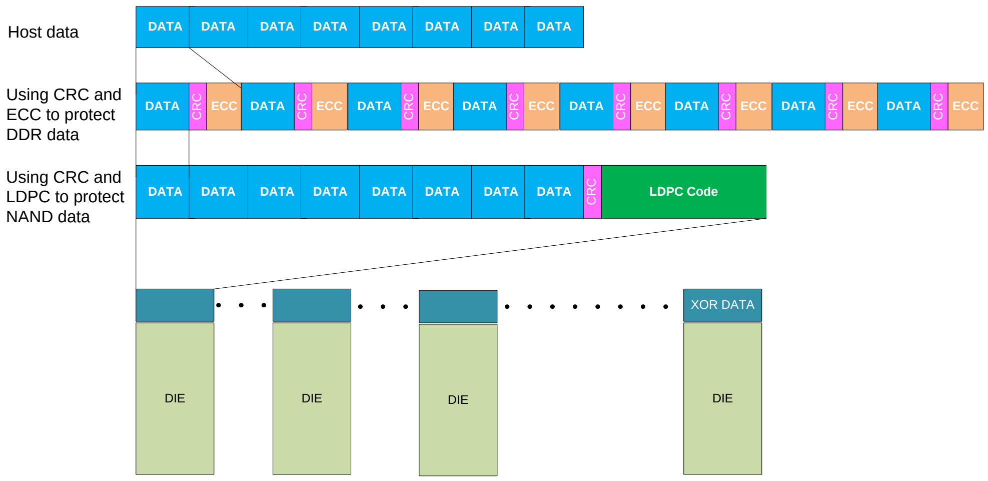
During the production of NAND flash chips, some storage units that do not comply with standards may be produced. Such storage units will be marked as bad blocks and will not be used. Bad blocks also occur during the use of SSDs. Huawei Enterprise SSD have own criteria to determine which are bad blocks. The criteria are set up based on plentiful experiment data and application scenarios. They include the NAND flash P/E count, error types, and error frequencies. If a block is determined bad, its data will be migrated to functional blocks to prevent data losses. During its life cycle, an SSD may have 1.5% of blocks bad. Huawei Enterprise SSD have reserved space inside to replace bad blocks, maintaining sufficient storage space and safeguarding data security.

3.3 Data Protection

3.3.1 Data Redundancy

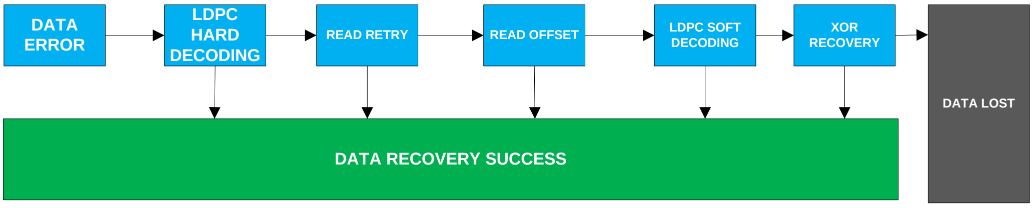
Huawei Enterprise SSD protect data all along the data path. On DDR memory, ECC and CRC are employed to prevent data errors and tampering due to exceptions on DDR memory. On NAND flash, LDPC and CRC are employed to prevent data losses due to exceptions on NAND flash. Among DIEs, XOR algorithm is employed to prevent data losses due to DIE or random failures.

Figure 3-4 Data protection diagram


Host data
DATA
Using CRC and ECC to protect DDR data
DATA
CRC
ECC
Using CRC and LDPC to protect NAND data
LDPC Code
XOR DATA
DIE

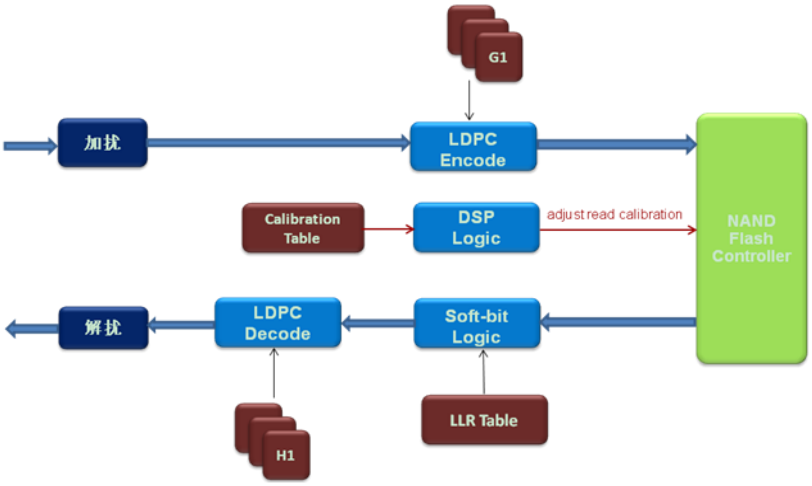
Huawei Enterprise SSD combine the LDPC algorithm with Read Retry and Read Offset technologies to maintain data reliability. When a host is writing data to NAND flash of an HUAWEI ENTERPRISE SSD, the HUAWEI ENTERPRISE SSD calculates the LDPC information of the host data and writes it together with the host data to NAND flash. When data is being read from NAND flash, data redundancy information is used to check data and correct errors if any. If an error occurs on NAND flash, the Huawei Enterprise SSD will perform LDPC hard decoding to rectify the error. If the rectification fails, the Huawei Enterprise SSD will perform other operations in sequence if needed until the error is rectified: read retry -> read offset > LDPC soft decoding > XOR recovery. See the data recovery process in the following figure:

Figure 3-5 Data protection diagram



Low-density parity-check (LDPC) codes are linear error correcting codes, which functionally are defined by a sparse parity check matrix. LDPC codes consist of encode, decode, soft-bit logic, and DSP logic. The H, G, LLR table, and calibration table are entries in the memory. H is the LDPC code parity check matrix and G is the LDPC code generation matrix. Multiple H and G items are options available to support flash chips from different manufacturers. The decoding process involves LDPC hard decoding and LDPC soft decoding. Compared with BCH, LDPC is theoretically more closer to Shannon Limit in terms of error correction capability. The LDPC hard decoding on Huawei Enterprise SSD offers a higher error correction capability than BCH under equivalent conditions. For example, on Micron 16 nm L95B cMLC, LDPC hard decoding offers an error correction capability four times as high as BCH.

Figure 3-6 LDPC diagram



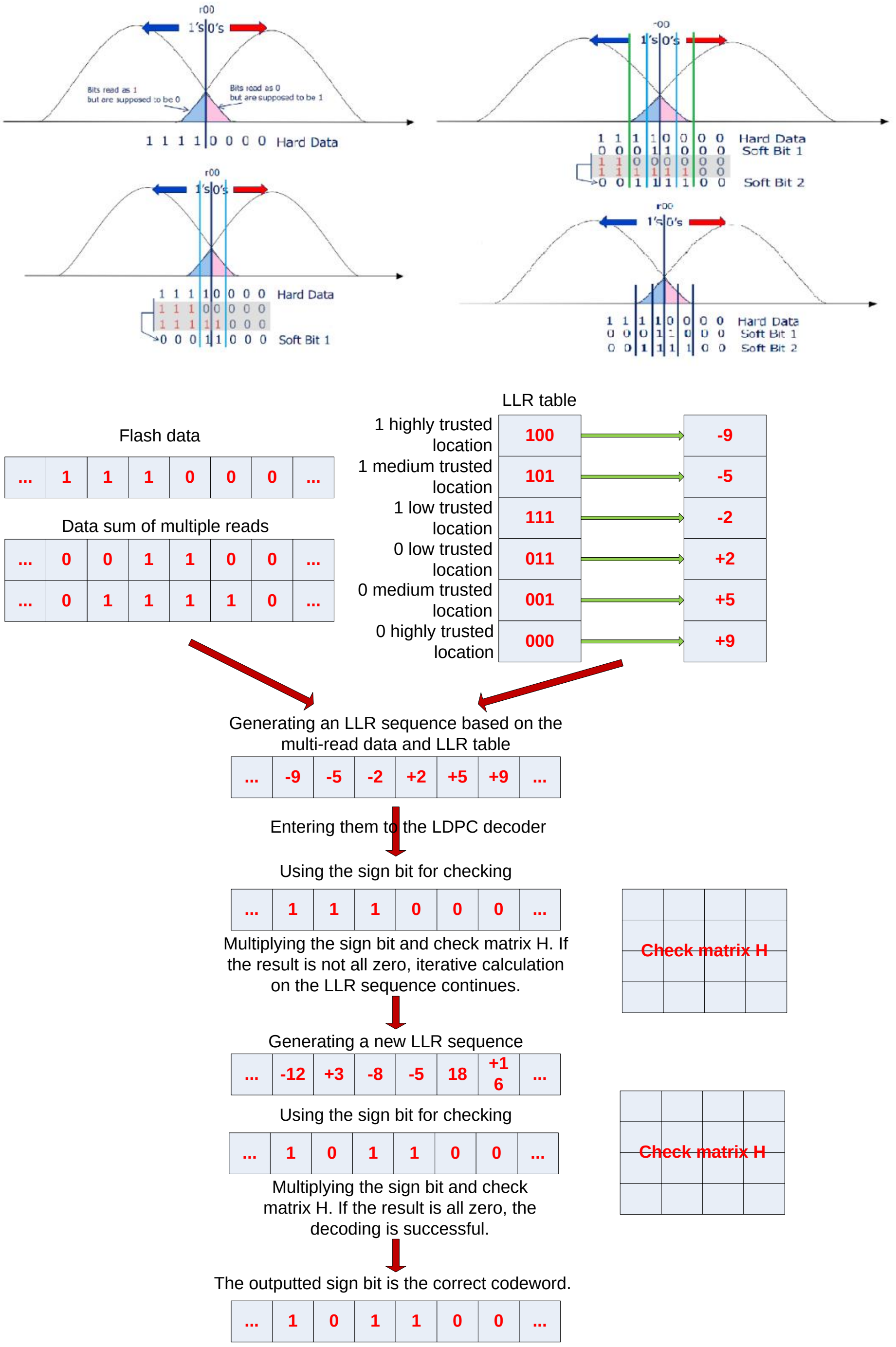
LDPC hard decoding uses the parity check matrix to check flash data in iterative mode. Each time after an iteration, the SSD controller multiplies the updated 0/1 matrix and the parity check matrix. If the result is all zero, the correct codeword is obtained and the iteration process is complete. Otherwise, data calibration is adjusted and the iteration process continues until the correct codeword is obtained or until the maximum allowed number of iterations is reached. Huawei Enterprise SSD adopt QC-LDPC codes and the specially designed two-step coding method to achieve a high coding speed while consuming fewer resources to store coding matrices. Benefiting from the optimized decoding concurrency with the layer NMSA algorithm, Huawei Enterprise SSD deliver an industry-leading performance: a throughput of 3.2 GB/s and a decoding latency of less than 3.1 μs.

Read Retry is a technology that recovers data by adjusting the flash chip read voltage. The read voltage levels are defined by the flash chip manufacturer. The firmware traverses those voltage levels to recover data. There are many read voltage modes for Read Retry. If the SSD controller retries each mode in sequence, the read latency is unacceptable. Huawei Enterprise SSD optimize Read Retry toward different application scenarios based on the block P/E count, retention duration, and read count. The most appropriate initial Read Retry mode is provided for each application scenario, reducing the required number of retries and increasing the read retry accuracy.

Read Retry offers only the manufacturer-defined read voltage levels, which do not fully meet actual requirements. Huawei Enterprise SSD add Read Offset to counteract the deficiencies of Read Retry. Huawei Enterprise SSD have a calibration table which is generated based on the manufacturer's private flash chip interfaces and a large number of samples and experiment results. The calibration table applies to different P/E counts, data retention duration, and read counts. It is saved on the Huawei Enterprise SSD. When an error occurs, the Huawei Enterprise SSD queries the table based on the block status (P/E count, data retention duration, and read count) to obtain the most appropriate read voltage, thereby increasing the success rate of data recovery.

LDPC soft decoding produces a real-number LLR sequence using the flash data and the preset data in the LLR table. Each real number specifies that the corresponding bit is the 0/1 possibility value. A positive number indicates a large possibility of 0, a larger absolute value representing a larger possibility of 0. A negative number indicates a large possibility of 1, a larger absolute value representing a larger possibility of 1. The LDPC decoder uses the LLR sequence to mark all the positions whose values are equal to or larger than 0 as 0 and all the positions whose values are smaller than 0 as 1. Then a 0/1 sequence is generated. If the result of the 0/1 sequence multiplied by the parity check matrix is all zero, the correct codeword is obtained and the iteration process is complete. Otherwise, the iteration process continues until the correct codeword is obtained or until the maximum allowed number of iterations is reached. Huawei Enterprise SSD optimize the LLR table to have a 150% higher error correction capability than hard decoding.

Figure 3-7 Working principle of soft decoding


r00
1's 0's
Bits read as 1
but are supposed to be 0
Bits read as 0
but are supposed to be 1
1 1 1 1 0 0 0 0 Hard Data
r00
1's 0's
1 1 1 1 0 0 0 0 Hard Data
1 1 1 0 0 0 0 0
1 1 1 1 1 0 0 0
0 0 0 1 1 0 0 0 Soft Bit 1
r00
1's 0's
1 1 1 1 0 0 0 0 Hard Data
0 0 0 1 1 0 0 0 Soft Bit 1
1 1 0 0 0 0 0 0
1 1 1 1 1 1 0 0
0 0 1 1 1 1 0 0 Soft Bit 2
r00
1's 0's
1 1 1 1 0 0 0 0 Hard Data
0 0 0 1 1 0 0 0 Soft Bit 1
0 0 1 1 1 1 0 0 Soft Bit 2

Flash data

...	1	1	1	0	0	0	...

Data sum of multiple reads

...	0	0	1	1	0	0	...
...	0	1	1	1	1	0	...

LLR table

1 highly trusted location	100	-9
1 medium trusted location	101	-5
1 low trusted location	111	-2
0 low trusted location	011	+2
0 medium trusted location	001	+5
0 highly trusted location	000	+9

Generating an LLR sequence based on the multi-read data and LLR table

...	-9	-5	-2	+2	+5	+9	...

Entering them to the LDPC decoder

Using the sign bit for checking

...	1	1	1	0	0	0	...

Check matrix H

Multiplying the sign bit and check matrix H. If the result is not all zero, iterative calculation on the LLR sequence continues.

Generating a new LLR sequence

...	-12	+3	-8	-5	18	+16	...

Using the sign bit for checking

...	1	0	1	1	0	0	...

Check matrix H

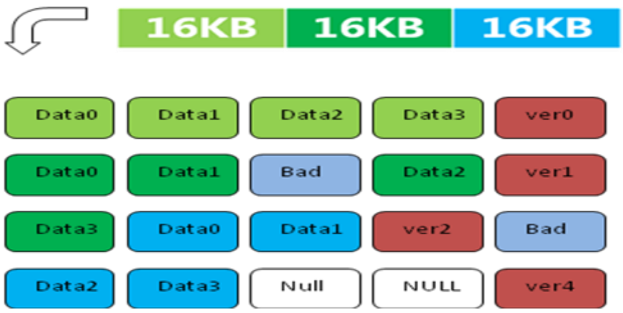
Multiplying the sign bit and check matrix H. If the result is all zero, the decoding is successful.

The outputted sign bit is the correct codeword.

...	1	0	1	1	0	0	...

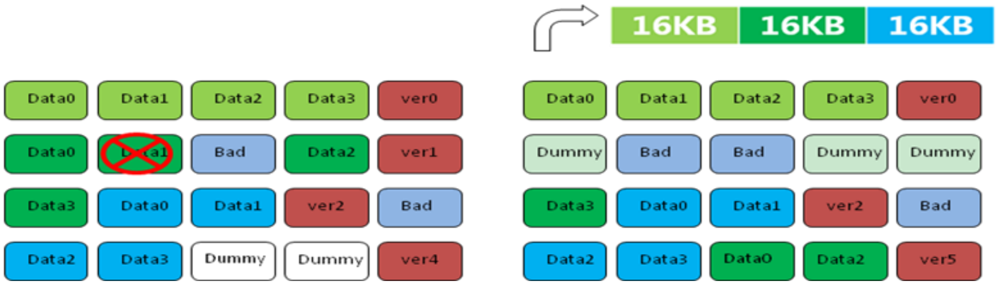
Huawei Enterprise SSD have an embedded XOR engine to provide data redundancy protection. When a physical fault (with pages, blocks, DIEs, or even the chip) occurs on the flash chip, Huawei Enterprise SSD use the check data blocks to recover the user data on faulty blocks, preventing user data losses. An Huawei Enterprise SSD has N+1 pages to form a stripe, and data is written to the flash in the unit of stripes. If a write encounters bad pages, those bad pages will be skipped and the write will be performed in a stripe of (N-1)+1 pages.

Figure 3-8 Writing I/O data to the flash


16KB 16KB 16KB
Data0 Data1 Data2 Data3 ver0
Data0 Data1 Bad Data2 ver1
Data3 Data0 Data1 ver2 Bad
Data2 Data3 Null NULL ver4

When an UNC error occurs on the flash chip, a data recovery operation is triggered. The remaining data on the faulty stripe is used in XOR calculation and the recovered data is written to an idle stripe. The faulty stripe will be reclaimed and converted into an (N-1)+1 stripe for later use.

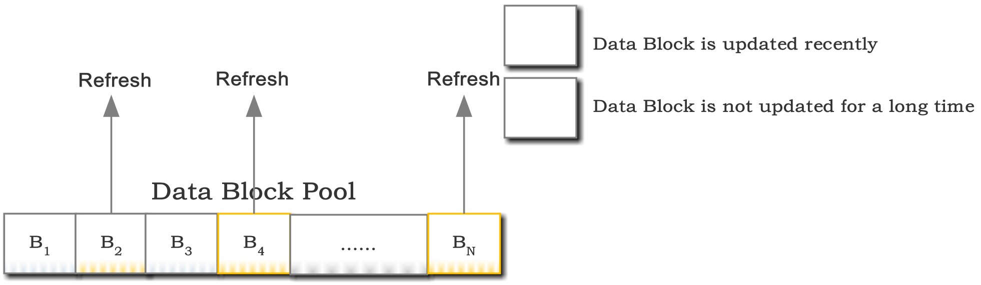
Figure 3-9 Recovering the read I/Os on bad blocks


16KB 16KB 16KB
Data0 Data1 Data2 Data3 ver0
Data0 Data1 Bad Data2 ver1
Data3 Data0 Data1 ver2 Bad
Data2 Data3 Dummy Dummy ver4
Data0 Data1 Data2 Data3 ver0
Dummy Bad Bad Dummy Dummy
Data3 Data0 Data1 ver2 Bad
Data2 Data3 Data0 Data2 ver5

3.3.2 Background Inspection

Data on NAND flash may have errors due to factors such as long-time retention, read interference, write interference, and random failures. Preventive inspection helps detect risks and imminent faults. Huawei Enterprise SSD combine read and write inspection to minimize data losses. Read inspection is conducted every two or three days. The SSD controller traverses the data on NAND flash to observe data errors. Those data with too many bit errors will be migrated in a timely manner. Read inspection helps prevent data losses caused by random failures and read interference. Write inspection is conducted every two weeks within normal temperature ranges (adjustable based on temperature). Write inspection checks the data retention duration and migrates those data whose data retention has exceeded the upper limit. Write inspection help prevents those errors caused by long-time data retention. In a word, background inspection detects data loss risks in a timely manner and prevents most of NAND flash errors, enhancing data reliability.

Figure 3-10 Working principle of background inspection


Data Block is updated recently
Data Block is not updated for a long time
Refresh
Refresh
Refresh
Data Block Pool
B1
B2
B3
B4
......
BN

3.3.3 Temperature Control

As the ambient temperature increases, the internal temperature inside an SSD increases also. When the internal temperature exceeds the threshold, the reliability and service life of components on the SSD are compromised. Each Huawei Enterprise SSD has two temperature sensors inside to monitor its internal temperature. If the internal temperature increases, the SSD controller decreases the performance level to reduce the internal temperature; if the internal temperature decreases, the SSD controller increases the performance level to make a better use of the Huawei Enterprise SSD.

Table 3-1 Huawei Enterprise SSD temperature and performance

Drive	Performance
Tj > 78 ℃	100% (High temperature alarm)
78 ℃ < Tj < 85 ℃	100% (Service life cannot be assured.)
Tj > 85 ℃, shorter than 6 minutes	50% (Major alarm)
Tj > 85 ℃, longer than 6 minutes	25%
Tj < 78 ℃	100% (Restrictions on performance are removed.)

3.3.4 Power Failure Protection

SSDs cache some data in DDR memory to improve performance. After a power failure occurs, backup power units on SSDs flush those data to NAND flash. The backup power design is a major factor affecting SSD reliability. The voltage monitoring module on SSDs monitors the input voltage in real time. When the voltage decreases to the preset value, the voltage monitoring module sends an interrupt instruction to the SSD controller, which then initiates a data flush process.

Huawei Enterprise SSD use solid aluminum capacitors that feature solid stability and reliability. Solid aluminum capacitors can work within the range of –50 ℃ to +105 ℃. Huawei Enterprise SSD have backup power units in groups. If a capacitor group malfunctions, another group can still sustain backup power. Huawei Enterprise SSD have 20% redundant backup power, maintaining normal backup power supply even when a single capacitor fails. In addition, Huawei Enterprise SSD periodically check the status of capacitors. If the backup power is unable to support data flushing, alarms will be generated and reported.

3.3.5 Data Encryption

To avoid data disclosure in disk abandoning, theft, or reuse and reduce the risks of information disclosure, some SSDs support data encryption.

Huawei Enterprise SSD complies with computer-related information security standards and protocols of Trusted Computing Group (TCG). Data at rest (DAR) encryption of Huawei Enterprise SSD prevents unauthorized users from accessing confidential enterprise data or sensitive user data when a disk is transferred, maintained, or recycled. Huawei Enterprise SSD use the AES256 encryption algorithm (key length is 256 bits) and support the XTS encryption mode commonly used in the industry to protect user data.

Trusted Computing Group

4 Customer Benefits

4.1 Constant Availability

Huawei is the only storage provider that possesses capabilities in researching and developing storage arrays, SSDs, and SSD controllers. With independent intellectual property rights in SSDs, Huawei has mastered core technologies in SSD controller chips. With a vertical integration capability in the solid state storage field, Huawei is able to provide service reliability protection solutions, convenient maintenance warranty programs, and customized end-to-end solutions, significantly improving Huawei Enterprise SSD availability.

- Intelligent bad block repair

  After a flash chip has been used for a long period of time, bad blocks (similar to bad sectors on HDDs) may appear on it. Huawei Enterprise SSD automatically detect bad blocks in the background. When this intelligent detection technology is used together with the intelligent bad block repair technology on Huawei self-developed storage arrays, Huawei Enterprise SSD report the UNC errors detected in the background to the host, helping recovering data in a timely manner and reducing the possibility of dual-SSD failure.

- Global wear leveling

  Huawei Enterprise SSD provide an internal wear query interface. A storage array can use this interface to obtain wear information about Huawei Enterprise SSD. In addition, an array-level wear policy can be made to extend the SSD service life and reduce the array rate. For example, a storage array implements global wear leveling to extend the service life of all SSDs and implements global anti-wear leveling to prevent a concurrent failure of multiple SSDs in a RAID group, thereby avoiding data loss. In doing so, the system availability is improved.

- Global FTL

  Huawei is the only in the industry to support full-stripe data flushing and dual controllers' concurrent writing to disks. Sequent I/O writes to Huawei Enterprise SSD enable Huawei Enterprise SSD to reduce the write amplification coefficient by 75%, prolonging the service life of Huawei Enterprise SSD.

- Non-disruptive upgrade

  Huawei Enterprise SSD and storage arrays combine to enable non-disruptive firmware upgrade. During an upgrade and activation process, storage arrays cache I/Os for Huawei Enterprise SSD to complete the upgrade and activation without affecting the customer's services.

- Data recovery

When data losses occur due to external factors or internal faults on SSDs, Huawei Enterprise SSD are capable of recovering most of data in a partial failure model.

- Quick data destruction

  Huawei Enterprise SSD support the T10 Sanitize feature, which destructs the data on an SSD within one or two minutes, saving time and cost for customers in data destruction.

4.2 High Performance

The advanced hardware architecture of Huawei Enterprise SSD employs powerful Cortex-A9 chips, low-power DDR4 memory, and up to 18 NAND flash channels. Additionally, Huawei Enterprise SSD adopt the hardware FTL architecture (unique in the industry), on which hardware acceleration along I/O paths makes multiple accelerators to coordinate for shorter I/O latency and higher throughput.

4.2.1 Performance Reference

Multiple application scenarios have sequential I/O models, for example, video applications and file copy. As shown in the two figures below, Huawei Enterprise SSD deliver the highest sequential read and write bandwidths in the industry.

Figure 4-1 Sequential read bandwidth

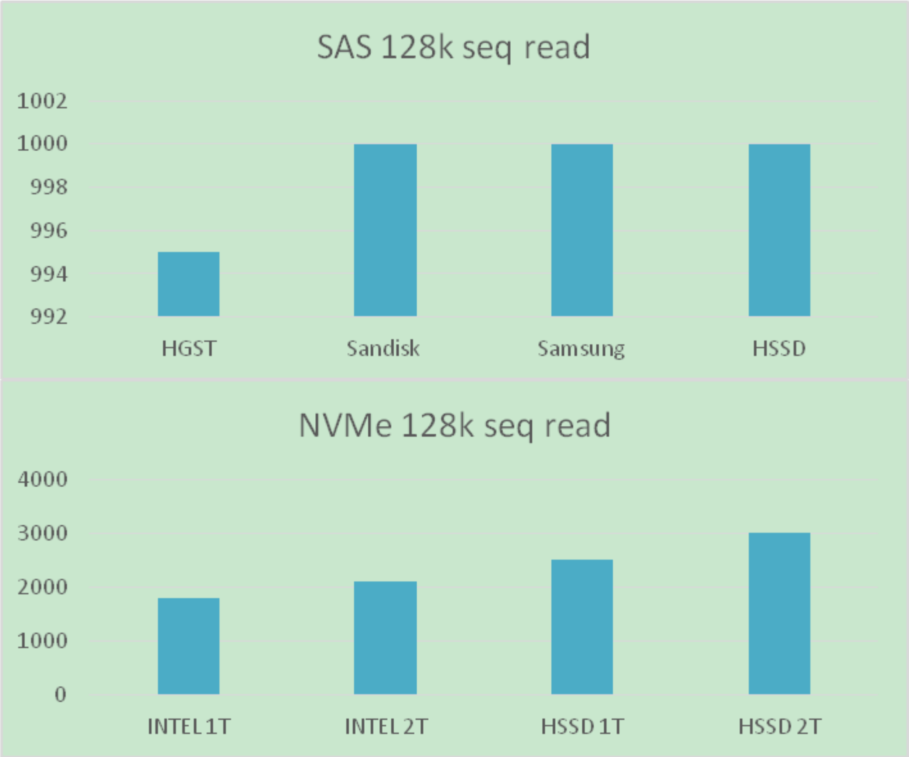
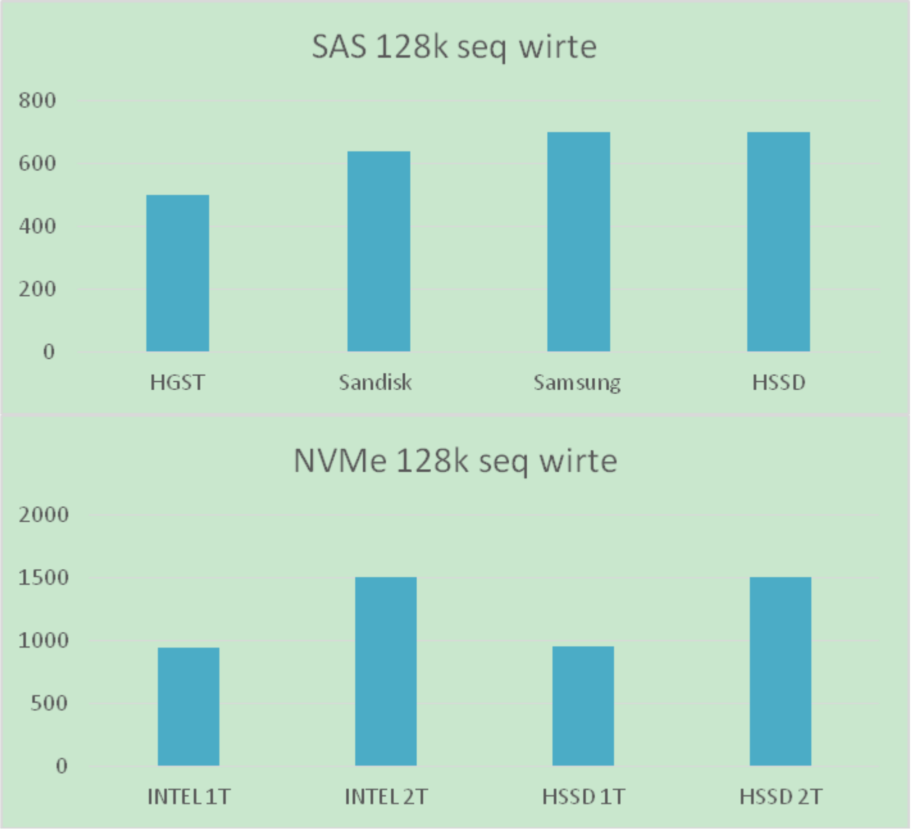


Figure 4-2 Sequential write bandwidth



Besides superior bandwidths, Huawei Enterprise SSD also deliver brilliant random access performance, and are a compelling choice for database and online transaction processing (OLTP) services. Huawei Enterprise SSD deliver the highest IOPS and the shortest latency under the same cost conditions in the industry.

Figure 4-3 Random read IOPS

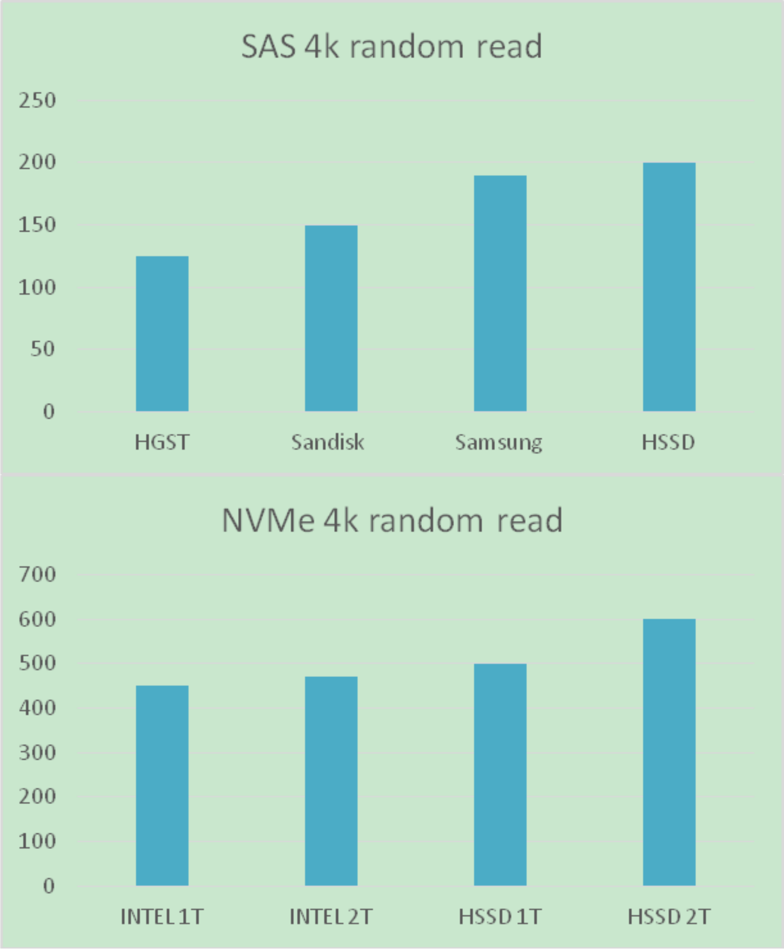


Figure 4-4 Random write IOPS

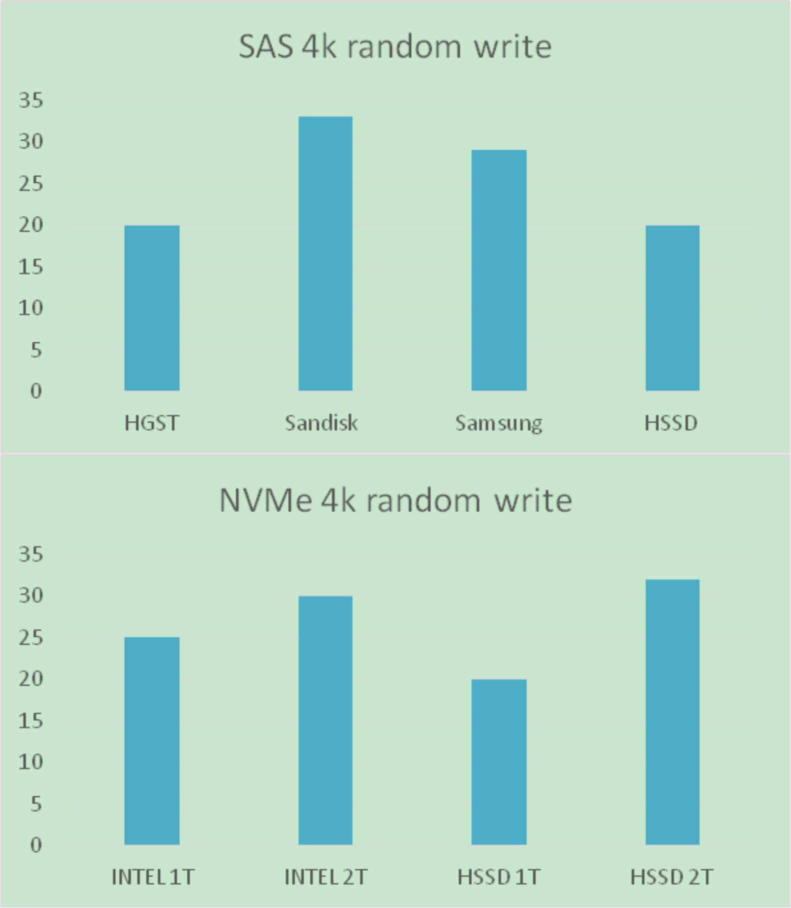


Figure 4-5 QD1 read latency

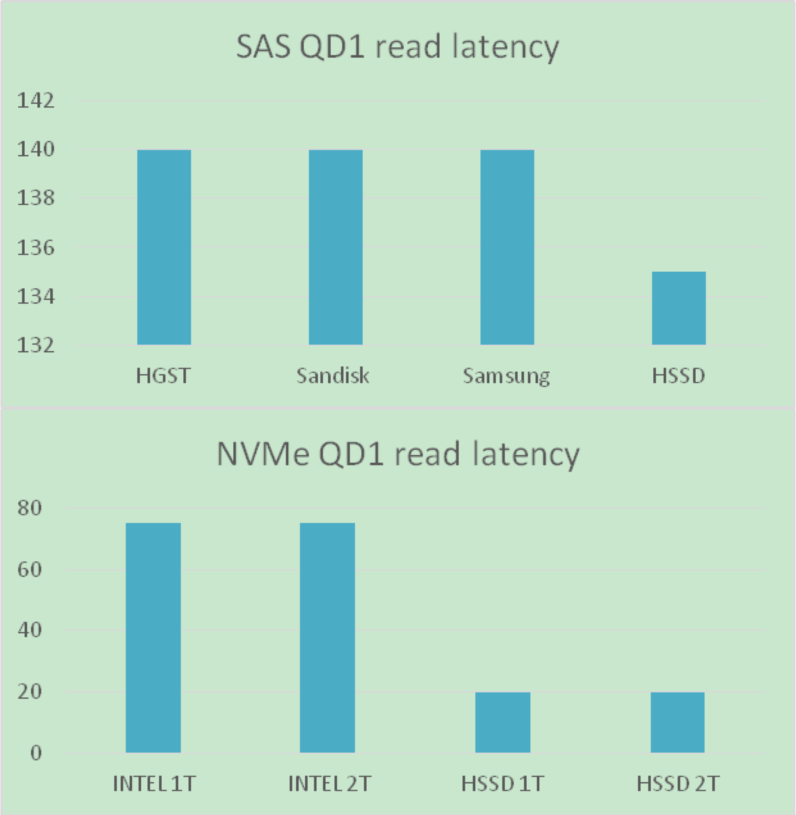
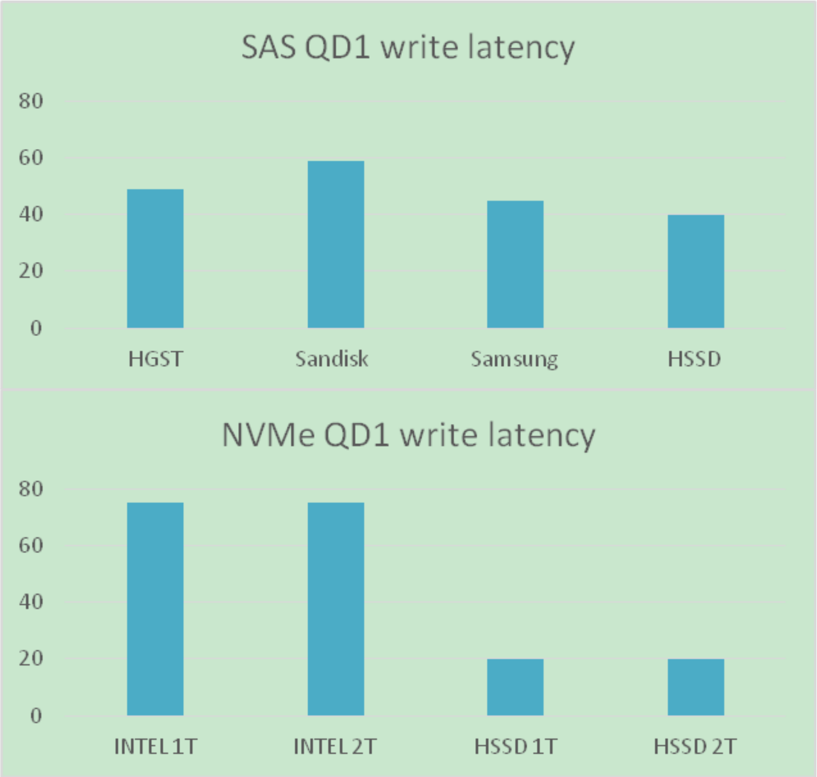


Figure 4-6 QD1 write latency



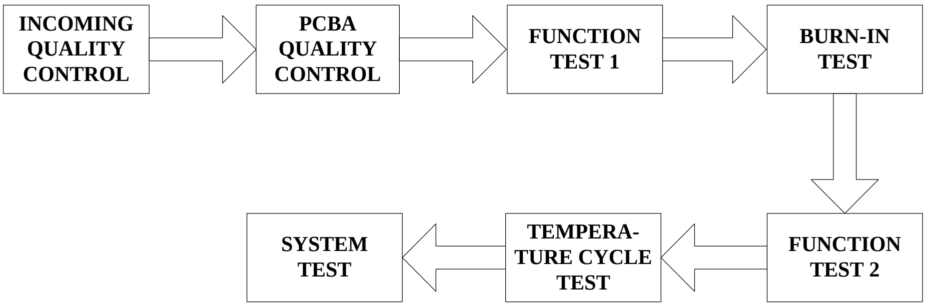
4.3 Robust Reliability

The committed mean time between failures (MTBF) of Huawei Enterprise SSD Diamond5 series is 3.0 million hours. Huawei takes the following measures to ensure robust reliability of Huawei Enterprise SSD.

4.3.1 Strict Filtering Check

The following figure shows the strict filtering process of Huawei Enterprise SSD. The process covers six major procedures and thousands of quality test items. It is worth mentioning that temperature cycle tests are applied to Huawei Enterprise SSD production for the first time, remarkably improving fault interception capabilities.

Figure 4-7 Huawei Enterprise SSD production filtering check diagram



- Incoming quality control

  Huawei strictly filters every incoming material of critical components such as NAND flash, DRAM, and backup power capacitors, and formulates strict quality acceptance standards. If the quality of Huawei Enterprise SSD does not meet the standards, Huawei will not sell them.

- Printed circuit board assembly (PCBA) quality control

  Automated optical inspection (AOI), automatic X-ray inspection (AXI), and in-circuit test (ICT) are applied on Huawei Enterprise SSD to eliminate defective components or the defects introduced during production.

- Function test 1

  Tests are conducted to check firmware, as well as verify the connectivity and basic functions of DDR, NOR flash, NAND flash, and capacitors.

- Burn-in test

  24-hour filtering check tests are conducted on DDR and NAND flash memory under high ambient temperature. Up to night check algorithms are used to filter out defective DDR components. Multiple rounds of read and write operations with different data models are performed to detect potentially faulty flash chips, ensuring the average outgoing quality (AOQ) of flash chips.

- Function test 2

  Six-hour I/O tests are conducted under normal ambient temperature, ensuring high reliability of Huawei Enterprise SSD.

- Temperature cycle test

  Within the –5 ℃ to +50 ℃ temperature cycle environment, eight-hour I/O read/write and power failure tests are conducted on Huawei Enterprise SSD to detect hardware and welding deficiencies.

- System test

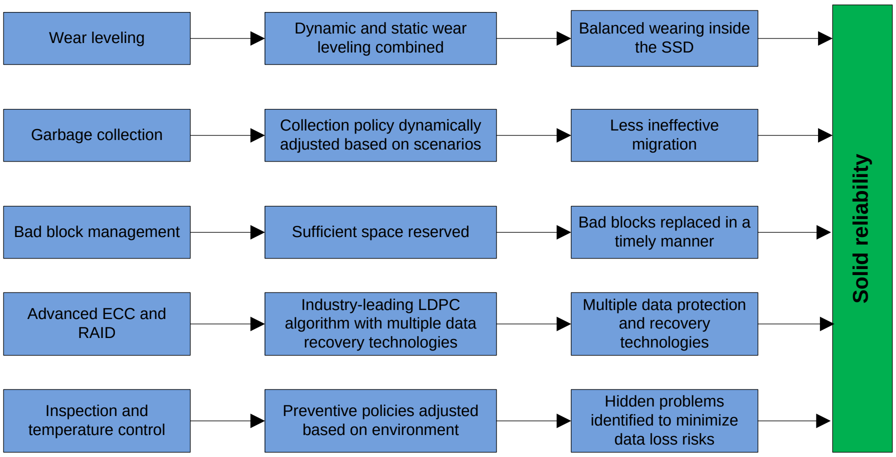
  After Huawei Enterprise SSD are installed on storage arrays, compatibility and basic function tests are conducted to verify the availability of Huawei Enterprise SSD.

Huawei strictly filters the hardware boards, welding processes, and functions of all Huawei Enterprise SSD and tests the working time and anti-pressure capabilities of Huawei Enterprise SSD in high temperature and temperature cycle environments. The Huawei Enterprise SSD that do not meet the requirements are filtered out and will not be sold.

4.3.2 Drive-Level Fault Tolerance

Huawei Enterprise SSD have an advanced software design, which is represented by the background inspection algorithm, bad block identification, DDR data protection, and all-path data protection. Additionally, the ECC error detection and correction algorithm and advanced power failure protection policies are applied to hardware of Huawei Enterprise SSD, ensuring data integrity and consistency.

Figure 4-8 Huawei Enterprise SSD life cycle diagram


Wear leveling
Dynamic and static wear leveling combined
Balanced wearing inside the SSD
Garbage collection
Collection policy dynamically adjusted based on scenarios
Less ineffective migration
Bad block management
Sufficient space reserved
Bad blocks replaced in a timely manner
Advanced ECC and RAID
Industry-leading LDPC algorithm with multiple data recovery technologies
Multiple data protection and recovery technologies
Inspection and temperature control
Preventive policies adjusted based on environment
Hidden problems identified to minimize data loss risks
Solid reliability

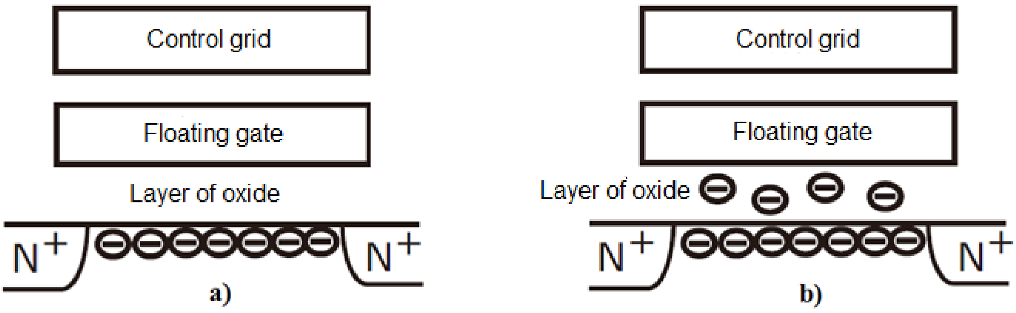
Apart from common reliability enhancement measures, Huawei improves reliability of Huawei Enterprise SSD based on the characteristics of internal key components and fault modes. After collecting the fault statistics about a large number of sample SSDs provided by Huawei and other vendors, Huawei locates the main cause of SSD failures: flash chip failure, which is reflected in too many UNCs and DIE failures. Huawei Enterprise SSD are capable of tolerating the preceding fault modes. In addition, Huawei Enterprise SSD take measures such as bad block scanning and health status monitoring to detect faults in advance and use array-level RAID and pre-copy technologies to ensure data security.

4.3.3 Trim Mode

A P/E period of NAND Flash is a round-trip movement of electron beams, which causes damage gradually to the layer of oxide and restricts the number of erasure times. How severely the layer of oxide is damaged depends on the P/E voltage stress exerted by NAND Flash chips upon the control grid. Before the delivery of NAND Flash chips, storage media vendors configure a series of special parameters (such as P/E voltage and time) on the chips to control the voltage stress exerted by NAND Flash chips upon the control grid during read, write, and erase operations. Reducing the voltage stress can increase the number of erasure times and prolong the service life of NAND Flash. The default parameter values are set to meet the specification requirements claimed by vendors and therefore may not be optimal choices for Huawei Enterprise SSD application scenarios. Accordingly, Huawei Enterprise SSD employs the Trim Mode technology. Based on the actual Huawei Enterprise SSD scenarios, Trim Mode customizes the special parameters on the chips and then adjusts the P/E voltage and time, increasing the number of chip erasure times, and prolonging the service life of NAND Flash.

Through the proprietary chip interfaces of vendors, Optimal Parameter Table based on different numbers of erasure times is obtained after many sampling experiments and saved in the Huawei Enterprise SSD. The service life of NAND Flash involves several phases. Huawei Enterprise SSD configure optimal parameters at different phases to ensure the best operating status and prolong the service lives of Micro 16 nm MLC chips by a maximum of 8 times.

Figure 4-9 NAND Flash storage units (a: Fresh NAND Flash; b: Cycled NAND Flash)



4.4 Long Service Life

The amount of data that is written into SSDs is limited because the erasure count of NAND flash is limited. However, the service life of SSDs is not equal to that of NAND flash. The service life of NAND flash is expressed by P/E cycle while the service life of SSDs is determined by the flash P/E, drive system design, and software algorithms.

4.4.1 SSD Service Life Evaluation

With the reference of generally used SSD lifetime evaluation method, Huawei calculates the lifetime of SSDs as follows:

$$Lifetime(Year) = \frac{TBW}{DWPD_{user} \times UserCapacity \times 365}$$

Where:

Terabytes Written (TBW): indicates the volume of data that can be written into an SSD in the life cycle.

Drive Write Per Day (DWPD$_{user}$): indicates the number of data writes on a disk per day. The DWPD here indicates the number of data writes on a disk per day at the site of the customer.

UserCapacity: indicates the visible capacity of an SSD.

365: indicates the total number of days in one year.

The claimed lifetime of enterprise-class SSDs is 5 years. However, the lifetime value is restricted according to different disk specifications. Generally, the DWPD is a factor that affects the lifetime. This factor is described in the specifications of SSDs provided by storage vendors. The claimed TBW of SSDs can be calculated as follows:

$$TBW = DWPD_{spec} \times 365 \times 5 \times UserCapacity$$

Where:

DWPD$_{spec}$: indicates the DWPD described in the specifications of SSDs provided by storage vendors.

UserCapacity: indicates the visible capacity of an SSD.

365: indicates the total number of days in one year.

5: indicates the lifetime of enterprise-class SSDs.

The formula used to calculate the lifetime can be simplified as follows:

$$Lifetime(Year) = \frac{DWPD_{spec} \times 5}{DWPD_{user}}$$

According to the preceding simplified formula, DWPD$_{spec}$ and DWPD$_{user}$ are necessary for calculating the lifetime of SSDs. Generally, the specifications of SSDs contain the claimed DWPD$_{spec}$ information about SSDs. The value of DWPD$_{user}$ must be calculated based on the service I/O model at the site. The formula is as follows:

$$DWPD_{user} = \frac{ThrouthputSpeed \times Write\% \times DutyCycle \times 3600 \times 24}{UserCapacity}$$

Where:

ThroughputSpeed: indicates the service throughput per second at the customer's site. The value is expressed in GB/s.

Write%: indicates the ratio of write service volume to the total service volume.

DutyCycle: indicates the ratio of the customer's service operating time to the total time.

UserCapacity: indicates the visible capacity of an SSD. The value is expressed in GB.

3600: indicates the number of seconds per hour.

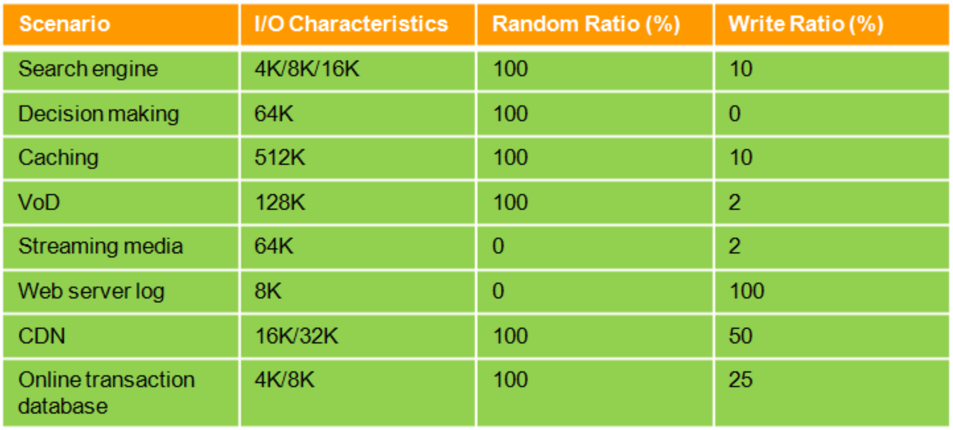
24: indicates the number of hours per day.

4.4.2 Huawei Enterprise SSD Lifetime Specifications

Specification list describes the lifetime specifications of Huawei Enterprise SSD Diamond5 (5-year warranty or the specified data write amount, whichever comes first).

The following table describes I/O characteristics under typical application scenarios:

Figure 4-10 I/O characteristics under typical application scenarios



Scenario	I/O Characteristics	Random Ratio (%)	Write Ratio (%)
Search engine	4K/8K/16K	100	10
Decision making	64K	100	0
Caching	512K	100	10
VoD	128K	100	2
Streaming media	64K	0	2
Web server log	8K	0	100
CDN	16K/32K	100	50
Online transaction database	4K/8K	100	25

Data source: Industry Standard Benchmarks, Customer Engagement Data

- Example 1

  An online translation database uses SSDs whose IOPS are 10,000 each. The throughput speed is 4k x 10,000 = 0.04 GB/s. The duty cycle is 50%. If the 400 GB Huawei Enterprise SSD Dimond5 is used:

  $$DWPD_{user} = \frac{0.04 \times 25\% \times 50\% \times 3600 \times 24}{400} = 1.08$$

  Therefore, the lifetime is:

  $$Lifetime(Year) = \frac{1.5 \times 5}{1.08} = 6.94 \quad \text{(years)}$$

- Example 2

  A VoD application uses SSDs whose throughput is 200 MB/s each. The throughput speed is 0.2 GB/s. The duty cycle is 100%. If the 600 GB Huawei Enterprise SSD Dimond5 is used:

  $$DWPD_{user} = \frac{0.2 \times 2\% \times 100\% \times 3600 \times 24}{600} = 0.576$$

  Therefore, the lifetime is:

  $$Lifetime(Year) = \frac{1.0 \times 5}{0.576} = 8.68 \quad \text{(years)}$$

4.5 Low TCO

SSDs are generally more expensive than HDDs, and some enterprises choose HDDs instead of SSDs due to tight budget. To help enterprises address this cost challenge, Huawei Enterprise SSD offer mainstream performance and reliability specifications as well as outstanding cost-effectiveness. Huawei Enterprise SSD products can replace 15k SAS HDDs at equivalent capacity and price, making Huawei go ahead of competitors six to twelve months.

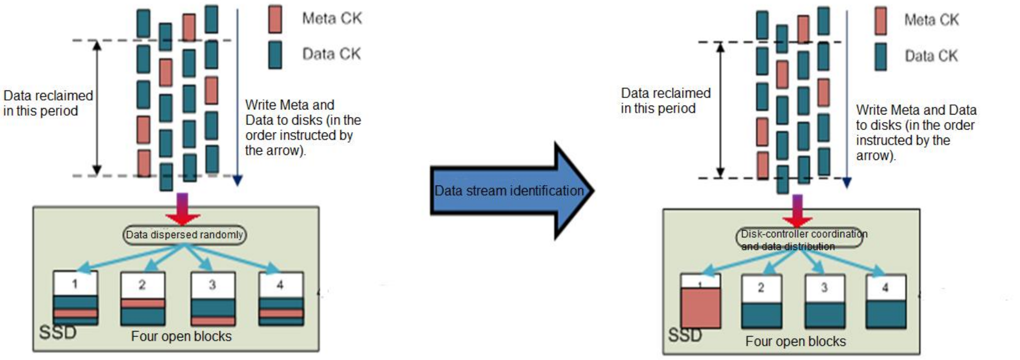
The cost competitiveness is further enhanced by the joint innovation between Huawei and partners. Huawei has signed Memorandum of Understanding (MoU) with major NAND flash vendors. These NAND flash products adopt cutting-edge manufacturing process and next-generation flash technology, making Huawei Enterprise SSD outshine all other competitors in price/performance ratio.

5 Advanced Features

5.1 Data Stream Identification

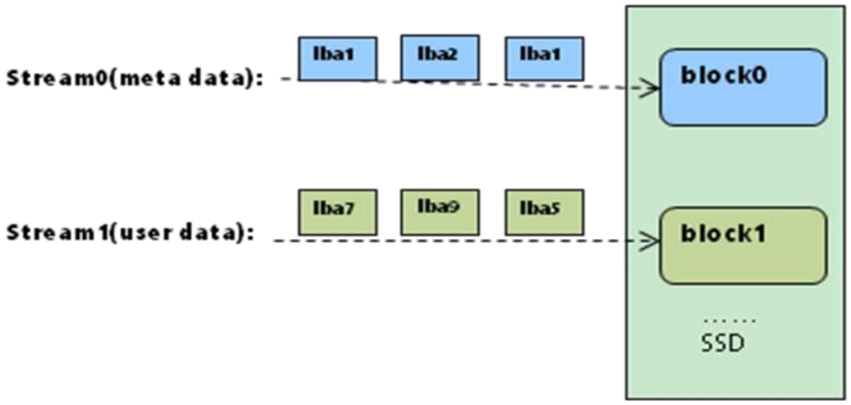
To improve the efficiency of SSD garbage collection, reduce SSD write amplification, and improving SSD write performance, vendors divide the data written to SSDs into multiple data streams based on the update frequency. SSDs identify these data streams and write the data of the same characteristics together to consecutive physical spaces on the disks.

Huawei Enterprise SSD work with Huawei storage arrays using FlashLink technology to optimize the data write and reclaiming processes. Only a small amount of data in SSDs needs to be migrated to release blocks completely, reducing write amplification, improving performance, and prolonging service lives of SSDs.

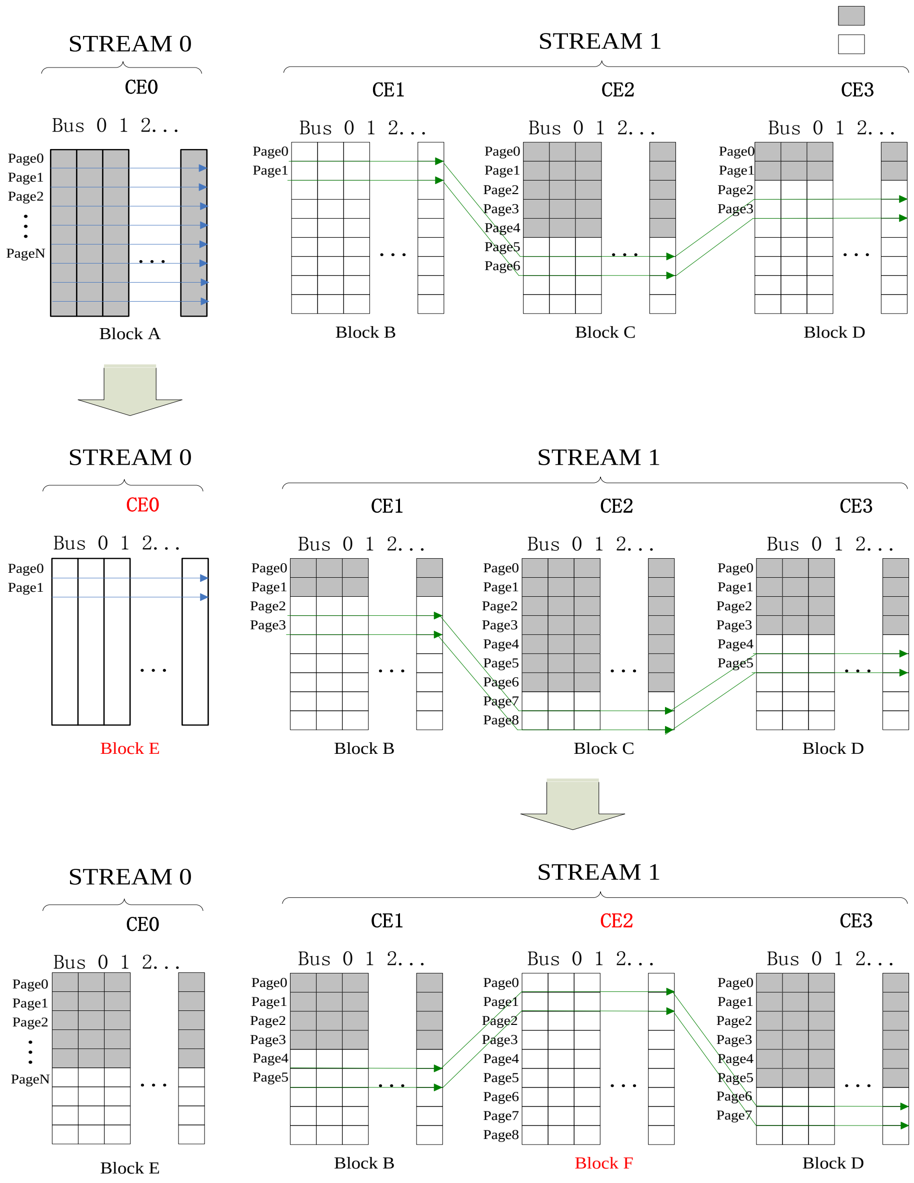

Meta CK
Data CK
Data reclaimed in this period
Write Meta and Data to disks (in the order instructed by the arrow).
Data dispersed randomly
1
2
3
4
SSD
Four open blocks
Data stream identification
Meta CK
Data CK
Data reclaimed in this period
Write Meta and Data to disks (in the order instructed by the arrow).
Disk-controller coordination and data distribution
1
2
3
4
SSD
Four open blocks

5.1.1 Data Stream Identification Write

Huawei storage arrays identify cold and hot data using FlashLink technology and put the data of the same access frequency into a logical zone. Different data is distinguished based on LBAs.



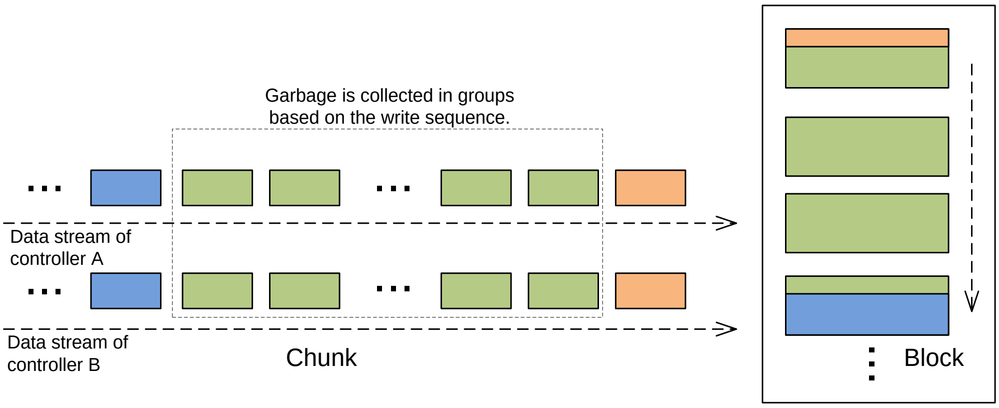
After receiving data of different LBAs, Huawei Enterprise SSD write the data in the same zone to the same block based on the zones preconfigured by the system. Huawei Enterprise SSD support multi-channel concurrency to ensure disk performance.



5.1.2 Global Garbage Collection

Huawei storage arrays work hand in hand with Huawei Enterprise SSD using FlashLink technology to provide the global garbage collection function. The storage system uses an innovative collection method that blocks are grouped based on write sequence, improving the write model of SSDs. In this way, global garbage collection and garbage collection within SSDs can interact with each other. The system analyzes the amount of garbage data in a chunk group (CKG) and data writing time, and then selects a target CKG. After collecting garbage data on the CKG, the system informs each SSD through the Trim/Unmap command that chunks in this GKG group are released. This process releases many chunks within a short time. For example, 256 CKGs are released at the same time with each SSD releasing 256 chunks, namely 256 MB. Each SSD then has one or more blocks available without data migration, greatly reducing write amplification.

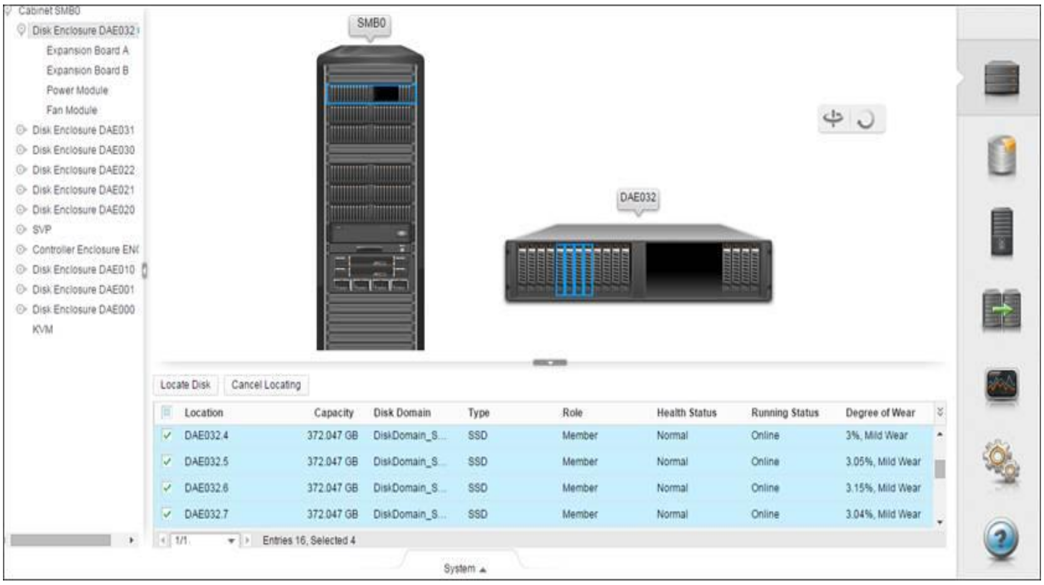
Figure 5-1 Global garbage collection


Garbage is collected in groups based on the write sequence.
...
...
...
...
Data stream of controller A
Data stream of controller B
Chunk
Block

In 0, the total capacity of the released chunks is triple that of blocks on an SSD. The SSD then has two blocks that can be erased directly (namely the two blocks in pure green). The two blocks with mixed colors also become erasable, if the system migrates data of one block because valid data in the two blocks with mixed colors occupies one block. The disk write amplification is (3+1)/3 = 1.3.

5.2 Service Life Prediction and Monitoring

Huawei storage arrays can monitor and predict the Huawei Enterprise SSD service life accurately, because the Huawei Enterprise SSD service life and the amount of written data can be queried. Therefore, Huawei can inform customer service of the possible disk faults in advance and notify customers to replace the disks one month before the service life expires.



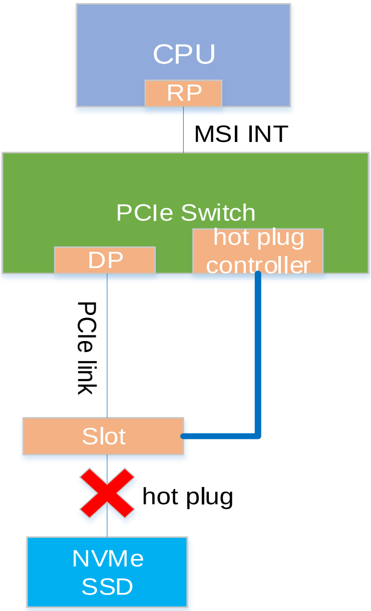
□	Location	Capacity	Disk Domain	Type	Role	Health Status	Running Status	Degree of Wear
✓	DAE032.4	372.047 GB	DiskDomain_S...	SSD	Member	Normal	Online	3%, Mild Wear
✓	DAE032.5	372.047 GB	DiskDomain_S...	SSD	Member	Normal	Online	3.05%, Mild Wear
✓	DAE032.6	372.047 GB	DiskDomain_S...	SSD	Member	Normal	Online	3.15%, Mild Wear
✓	DAE032.7	372.047 GB	DiskDomain_S...	SSD	Member	Normal	Online	3.04%, Mild Wear

1/1 Entries 16, Selected 4

5.3 NVMe Forcible Hot Swap

The NVMe SSD developed by Huawei supports NVMe 1.2 protocol and has dual ports. Two PCIe 3.0x2 links are independent from each other, providing hardware conditions for system restoration and troubleshooting, achieving redundancy protection, and improving the system reliability. With years of experience in PCIe links, Huawei has overcome the difficulty on PCIe link troubleshooting and developed the forcible hot swap feature for Huawei NVMe SSD and storage arrays.

As shown in the following figure, NVMe SSD hot swap causes an abnormal process to PCIe links. In Huawei storage arrays, the SSDs can be removed at any time in any manner, which achieves industry-leading NVMe productization.



A Appendix: SSD Terminology

WAF

Write amplification factor represents the amount of data NAND flash has to write in relationship to the amount of data the host's flash controller has to write. Without data compression, a factor of 1 is ideal. However, the factor is usually larger than 1 because NAND flash has to be erased before data is written in. The factor is affected by an SSD's over-provisioning space size and the GC algorithm. The larger the factor, the shorter the SSD's life and the lower its performance.

OP

Over-provisioning (OP), the practice of allocating a specific, permanent amount of free space on an SSD. For example, the capacity visible to users of an SSD is 400 GB but the total capacity of the SSD is 512 GB, then 112 GB has been over-provisioned or the over-provisioning space is 112 GB. The greater the proportion of over-provisioning space to NAND flash's total space, the greater the SSD's WAF.

FTL

The flash translation layer (FTL) is a software layer allowing an SSD to simulate HDD operations. Because NAND flash cannot be written in bytes and must be erased before programming, FTL main functions include organizing and managing host data, distributing the data onto NAND flash chips in an orderly manner. Its functions also include maintaining the mappings between logical block addresses (LBAs) and physical block addresses (PBAs), collecting garbage, leveling wear, and managing bad blocks.

GC

GC stands for garbage collection. NAND flash does not allow re-programming. As a result, when a host changes data on the same logical block, the SSD writes new data into different flash addresses and marks the original data as invalid. Multiple data modifications and updates cause a large number of blocks to contain valid pages and invalid pages at the same time, which means there are no empty blocks for new data. GC migrates data from multiple blocks to a new block and then erase the blocks whose valid data has been migrated so that they are available to new data.

Wear Leveling

NAND flash has a limited number of erasure times for blocks. If a block has run out of its erasure times, it is no longer usable. GC brings more wear. Generally, if an SSD has a large over-provisioning space, the wear brought by GC is small. If the flash life of some part of the over-provisioning space runs out, the over-provisioning space gets smaller during use, meaning that GC brings more wear. Therefore, the best-case scenario would be all flash chips have the same level of wear. Then they can reach their service lives at almost the same time. Wear leveling evenly distributes flash erasures to all chips, and the algorithm used for this is called wear leveling algorithm.

Bad Block

Each NAND flash module allows around 3% bad blocks during its lifecycle. Because bad blocks have non-ensured stability and reliability, they are avoided during SSD operation. The concept of bad block management then is introduced.

There are two types of bad blocks: factory bad blocks and initial bad/invalid bad blocks. Factory bad blocks are produced before the SSD is delivered. Therefore, they are marked by the vendor before shipment. Initial bad/invalid blocks are also called worn-out bad blocks and are produced during use. If erasure failure is reported for a block, it is identified as an initial bad block.

Power Loss Protection

Enterprise-level storage systems require robust reliability from power supply systems. However, power systems occasionally encounter faults and sometimes lead to power failure. Enterprise-level SSDs store mission-critical data and do not allow loss of data, be it stored or not. Therefore, SSD power system design must consider the data reliability in accidental power failures. Abrupt power failure during NAND flash programming contains data loss risks, especially for MLC. To prevent data loss caused by power failure during data writing, capacitors are used to flush the unwritten data, which is called SSD power loss protection.

DPP/DIF

DPP stands for data path protection and DIF for data integrity field. DPP protects the integrity and correctness of data in I/O paths. Besides ECC- and CRC-based validation for critical memory modules (such as DDR), SSDs also provides LBA check and DIF protection. In DIF, several bytes of protection information are added to the end of a user data sector to protect the data consistency between a host and the SSD.

TRIM

TRIM refers to a rapid data deletion method. A host delivers the TRIM command informing an SSD that the data within an LBA segment is no longer useful, then the SSD rapidly releases the flash space occupied by the data. TRIM is rapid because NAND flash space can be released by simply data migration. TRIM is a lot lower in the amount of data migrated than GC.

SSD

SSDs use semiconductor transistors (flash chips) instead of magnetic media as basic storage units to store data. Electronic reading and writing operations are used to replace motor rotation and mechanical arm addressing, which greatly reduces access latency and improves I/O access efficiency.

Huawei Enterprise SSD

Huawei SSD

DWPD

Drive writes per day (DWPD) tells how many times you can overwrite the entire capacity of the SSD every day of its usable life (5 years) without failure during the warranty period.

UBER

Uncorrectable bit error rate refers to the probability that an uncorrectable ECC error occurs, that is, there are more bits inverted within a codeword than can be corrected by the ECC algorithm. The UBER value reflects the probability of ECC errors in a codeword.