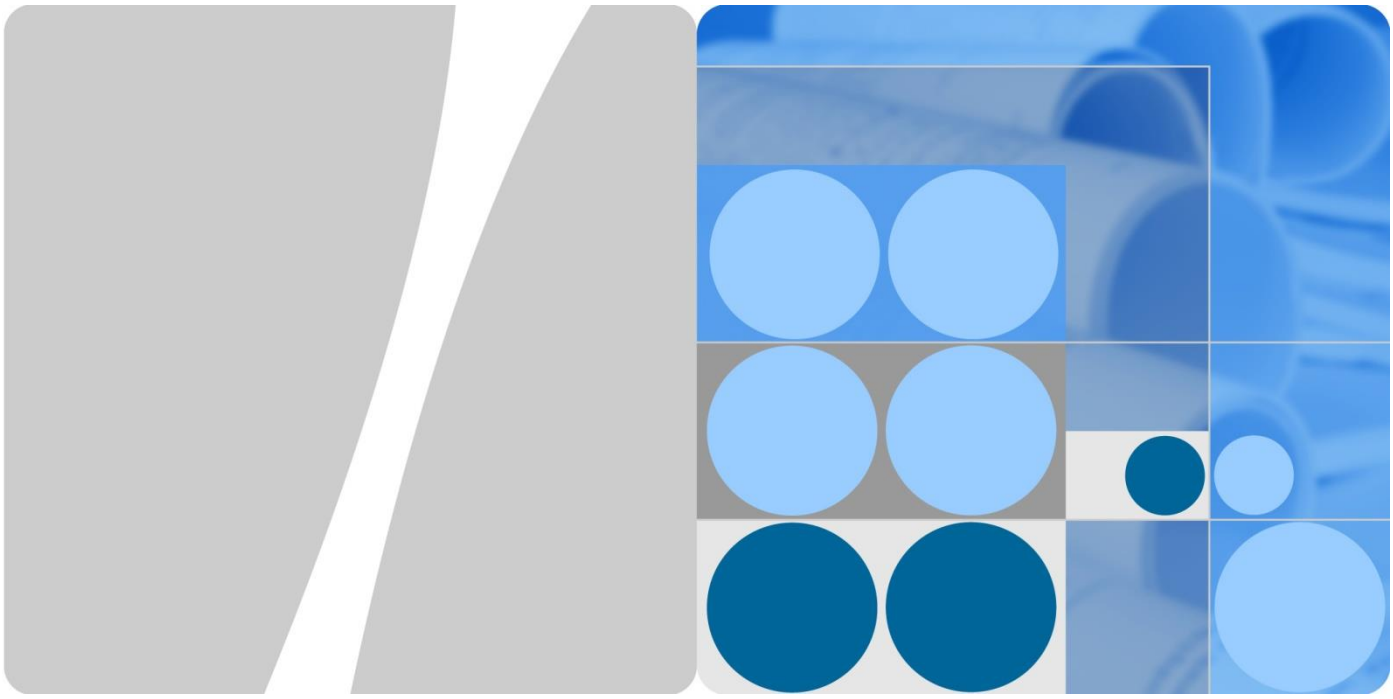


Huawei OceanStor 2600 V3 for Video Technical White Paper

Issue 01

Date 2017-01-17

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://e.huawei.com

Email: support@huawei.com

Contents

1 Abstract

As social economy develops rapidly, a series of problems involving social security, traffic, law enforcement in cities, and environment are arising. Related departments are required to speed up response to these problems and improve management efficiency to facilitate economic development. Video surveillance is applied to various industries, including:

- Traffic surveillance
- Urban safety surveillance
- Unattended area surveillance
- Mobile surveillance

With the development of video surveillance systems, the number of high definition (HD) cameras is rapidly increasing. Meanwhile, image quality has been developed from Common Intermediate Format (CIF) to 720p, 1080p, and even 4K. With explosive growth of surveillance data, dedicated storage systems are appreciated to meet high capacity and security requirements.

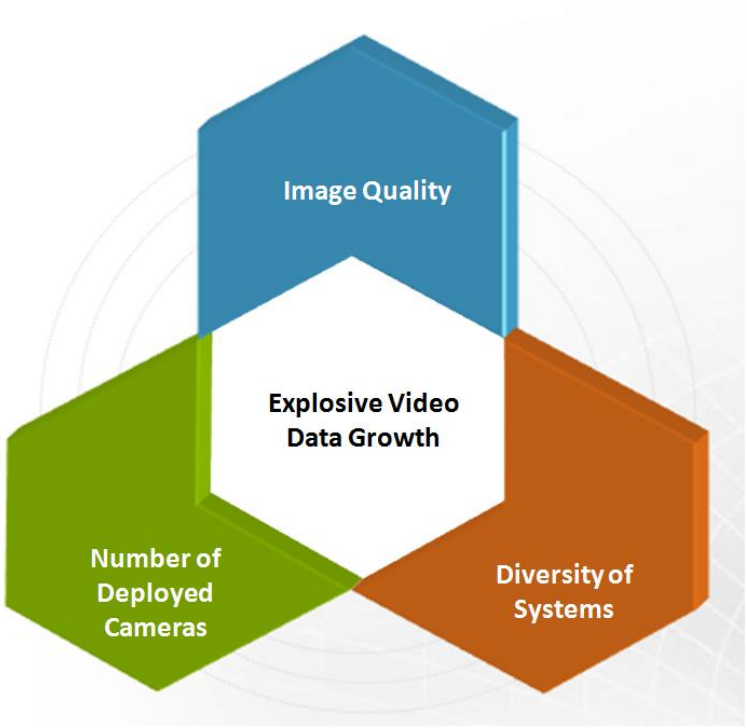
This document introduces video surveillance storage products in the industry as well as of Huawei and describes the advantages of Huawei products over those in the industry.

The advantage of Huawei video surveillance storage products can be clearly seen in this document, namely, being able to help customers meet needs of various video surveillance applications at a lower price and with high performance.

2 Overview

2.1 Development Trends of Video Surveillance Systems

2.1.1 Explosive Data Growth



- Image quality

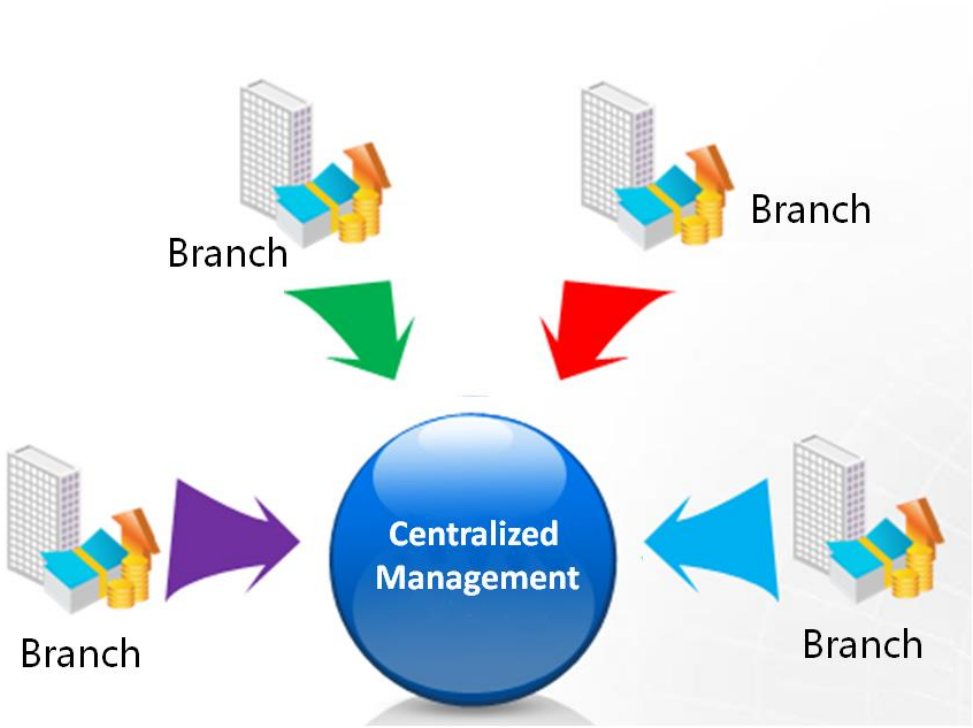
  The number of HD cameras is increasing. The resolution of cameras evolves from low definition (LD) D1 (2 Mbit/s) to HD 720p (4 Mbit/s), full HD 1080p (6 or 8 Mbit/s), and 4K (8 or 10 Mbit/s). The storage capacity is increased by three to five times. HD and full HD cameras have become the mainstream. 4K cameras are also used for some important surveillance scenarios.
- Number of deployed cameras

Video surveillance systems are widely used in fields such as urban traffic, safe cities, and unattended areas. The number of deployed cameras and capacities are also increasing exponentially.

- Diversity of systems

  Video surveillance systems are interconnected with access control systems, alarming systems, and other systems. Storage capacities are increasing for meeting data backup and archiving needs.

2.1.2 Popular Centralized Data Storage and Management



As video camera sampling points and capacities increase, management of video surveillance systems becomes more and more complex and maintenance costs go up, which require an increasing number of maintenance professionals. Therefore, centralized management is gaining popularity. Centralized management mode boasts the following benefits:

- Effectively reduces maintenance complexity and number of maintenance professionals as well as total cost of ownership (TCO).
- Improves maintenance efficiency and simplifies maintenance procedures.

2.1.3 Rapid Development of Large-Capacity Storage Systems

Due to the increase of camera sampling points and video surveillance image resolution and the need for centralized storage, the demand for large-capacity storage systems for video surveillance is increasing. To satisfy this demand, major storage vendors have all released large-capacity storage systems of one PB or above.

3 OceanStor 2600 V3 for Video

3.1 Introduction to OceanStor 2600 V3 for Video

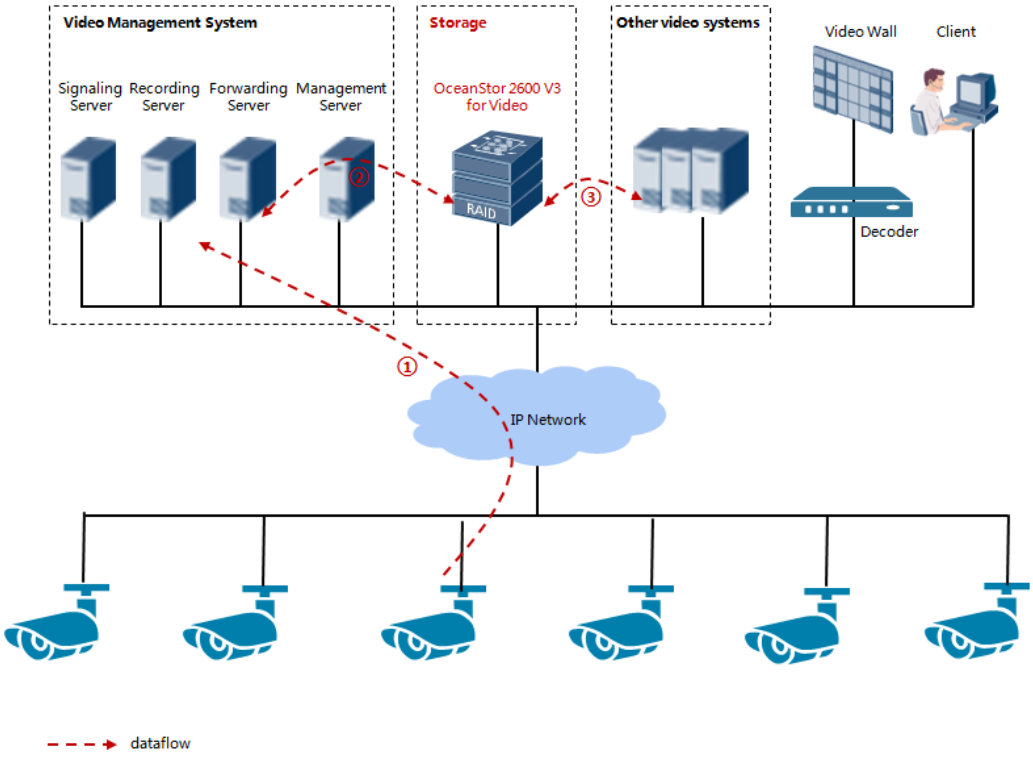
3.1.1 Overview

Huawei OceanStor 2600 V3 Video is the new-generation storage array dedicated for video surveillance. Based on the new-generation Huawei OceanStor V3 storage array, this edition is optimized for video surveillance scenarios to provide enterprise-level reliability and meet special requirements of IP video surveillance.

The storage system has dual controllers working in active-active mode, employs multi-core processors, supports 8/16 Gbit/s Fibre Channel and 10 Gbit/s FCoE host interfaces, adopts the new PCIe 3.0 bus technology, and uses 12 Gbit/s SAS 3.0 high-speed disk enclosure interfaces. The specifications and performance beat products of the same level in every aspect and can meet the requirements of centralized video surveillance storage for enhanced capacity and performance.

3.1.2 Typical Networking

The storage system supports four to 396 large-capacity disks and is applicable to video surveillance storage scenarios of different sizes. This following figure shows the typical network diagram.



Solution highlights:

- With innovative design, self-developed software and hardware, and controllable technologies, the storage system provides enterprise-level performance and reliability.
- The storage system has dual controllers working in active-active mode. Dual controllers provide services simultaneously. If one controller is faulty, the other controller takes over its services, reducing the risks of data loss and service interruption due to device faults.
- The storage system is highly scalable with flexible configurations, and can be used in large, medium, and small-sized video surveillance scenarios.
- Data is stored and managed centrally, reducing data management risks and cutting the TCO.
- The innovative Huawei SmartConfig tool is supported, freeing IT administrators from complex storage technologies and greatly simplifying the configuration process.
- Unified management: The storage system provides powerful storage management software that supports global topology view, capacity analysis, performance analysis, fault diagnosis, and end-to-end service visualization to manage a wide range of devices.

3.2 Huawei Video Surveillance Storage Solution

3.2.1 RAID 2.0+

Huawei RAID 2.0+ technology adopts a two-layer virtualization management model that consists of underlying disk management and upper-layer resource management. In a storage system, the space of each disk is divided into data blocks. RAID groups are created based on data blocks so that data can be evenly distributed on all disks in a storage pool. In addition, resources are managed based on data blocks, greatly improving the resource management efficiency.

Compared with traditional RAID groups, RAID 2.0+ has the following advantages:

- Load balancing is achieved among all disks in a disk pool, reducing the overall failure rate. RAID 2.0+ implements block virtualization to enable data to be automatically and evenly distributed onto all disks in a storage pool, preventing unbalanced loads. This approach decreases the overall failure rate of a storage system.
- Fast thin reconstruction significantly reduces the possibility of dual-disk failures. RAID 2.0+ breaks the performance bottlenecks seen in target disks (hot spare disks) that are used by traditional RAID groups for data reconstruction. As a result, the write bandwidth of reconstruction data flows is no longer a bottleneck for the reconstruction speed, greatly accelerating data reconstruction, decreasing dual-disk failure possibility, and improving storage system reliability.
- Self-check and self-healing mechanisms ensure system reliability. Based on hot spare policies, RAID 2.0+ automatically reserves specific hot spare space in disk domains. When the system automatically detects irrecoverable media errors in a specified area on disks or the whole disk fails, the system automatically reconstructs affected data blocks to hot spare space on other disks, achieving fast self-healing.

3.2.2 Reliability Technologies

I. System reliability: active-active dual-controller architecture

The storage system uses active-active dual-controller system architecture. Both controllers concurrently store and process video data from hosts, improving video surveillance performance and ensuring service continuity upon a controller failure.

II. System reliability: cache mirroring

The storage system supports setting of the cache mirroring function for services upon service deployment. After the cache mirroring function is configured, all the data written into the cache of the local controller is mirrored to the cache of the remote controller through dedicated high-speed PCIe channels. In this manner, dirty data of the local controller is saved in the cache of the remote controller. After data in the cache of the local controller is written to disks, data in the cache of the remote controller is deleted through dedicated high-speed PCIe channels. In this way, dirty data (data not written to disks) in the system is saved in caches of both controllers, namely, the dirty data in the system is backed up mutually.

If one controller fails, cache of the other controller will save dirty data of the failed controller. After the failed controller recovers, the dirty data in the cache of the remote controller is flushed into the disks of the local controller, ensuring the integrity of video surveillance data. No video surveillance data is lost upon a controller failure.

III. System reliability: RAID 6

As NL-SAS and SATA disk capacity increases, RAID group reconstruction period prolongs, and the probability of a concurrent failure of two disks grows. RAID 6 adopts double-disk parity and can tolerate a concurrent failure of two disks, greatly improving disk reliability. For a twelve-disk pool with nine RAID 5 disks, the annual probability of data loss is 0.051,560%, while for a twelve-disk pool with ten RAID 6 disks, the probability will fall to 0.000,026%, 1,995 times less than that of nine RAID 5 disks. The performance of SATA disks configured with RAID 6 is basically the same as those configured with RAID 5.

In general, although the disk utilization is relatively low when SATA disks are used to form a RAID 6 group, the system reliability is improved remarkably (the annual failure rate is reduced by 1,995 times).

IV. Disk reliability: disk pre-copy

The system periodically checks the health status of disks, locates the disks that are about to fail, and starts data reconstruction in advance. Compared with the disk pre-copy technology of traditional RAID groups, RAID 2.0+ uses the reconstruction mechanism to generate the data to be copied, further mitigating the performance pressure on a single disk imposed by pre-copying. The pre-copy technology eliminates the risks of both RAID degradation and data loss.

V. Disk reliability: bad sector repair

If errors caused by bad sectors occur during data read from disks, the bad sector repair technology can be used to repair the bad sectors.

Bad sector repair can repair as many bad sectors as possible, reducing the disk failure rate by over 50%. This extends the service life of disks and ensures the security of video surveillance data.

VI. Disk reliability: background scanning

The background scanning technology periodically checks disk health status and scans disks for bad sectors, to efficiently locate bad sectors that occur during disk running. The discovered bad sectors can be repaired by the bad sector repair technology.

In addition, the background scanning technology discovers hidden bad sectors, preventing performance from being degraded by errors and follow-up repair during service interactions. Meanwhile, the technology reduces RAID group failures due to bad sectors occurring on multiple disks in a disk pool all at once.

VII. Disk reliability: online diagnosis

The online diagnosis technology enables failed disks to be recovered without interrupting ongoing services. When a disk is faulty, or read or write errors occur in a disk pool:

1. If forcing the failed disk to go offline does not cause the RAID group to fail:

   Isolate the failed disk, retry the failed disk in the background, and power it off and on to recover it. After the failed disk is recovered, connect it to the system and start thin reconstruction to recover the RAID group quickly.

2. If isolating the failed disk will cause the RAID group to fail:

   Temporarily suspend data reads and writes on the failed disk, and try to recover it by retrying it or powering it on and off.

The preceding two methods can recover a failed disk without interrupting ongoing services.

The online diagnosis technology can be used to clear disk faults caused by non-hardware problems such as hardware firmware bugs, disk logic faults, and link logic errors. The online diagnosis technology reduces disk failure rates and improves service reliability.

VIII. Disk reliability: DHA

Disk health analyzer (DHA) is a technology used to analyze and evaluate disk health status online. This technology can establish a comprehensive system to collect and analyze disk health data and evaluate disk health status.

DHA periodically collects key disk indicators such as I/O models and S.M.A.R.T information. After collecting the data, DHA analyzes the data using a mathematical model and evaluates the current disk health status based on the analysis results to predict potential disk failures.

After DHA detects potential disk failures, it starts the pre-failure mechanism and uses the pre-copy technology to prevent disk faults. This reduces the RAID degradation probability and the risks of data loss and service interruption.

3.2.3 Image Repair and Video Recovery Technologies

I. Bad sector repair technology

The bad sector repair technology is an effective method in reducing disk failure rates and video and image data losses. Two diagnosis methods are available.

1. Automatic analysis upon read/write failures When a read or write failure occurs, the storage system analyzes the failure using the information such as current system status, current disk status, and I/O failure information and then determines whether the failure is caused by bad sectors. If the failure is indeed caused by bad sectors, the system automatically starts recovery.

2. Proactive disk media scanning. The storage system supports background scanning of disk media. In off-peak hours, the disk media is periodically checked to efficiently locate bad sectors and avoid leaving faults unaddressed. To prevent disk scanning from consuming back-end bandwidth, the storage system directly uses the built-in disk media scanning function rather than sequentially reads and writes all the sectors of a disk, minimizing the impact on system performance. If the disk is being accessed during scanning, the scanning automatically stops for the disk to continue to process read and write requests. After the read or write requests are complete, the scanning resumes from where it stopped.

With the RAID 2.0+ technology, bad sectors are automatically marked. After that, data on the bad sectors are automatically recovered using RAID algorithms, and the reconstructed data is rewritten into a selected hot spare space.

II. Video Image Repair (SecureVideo)

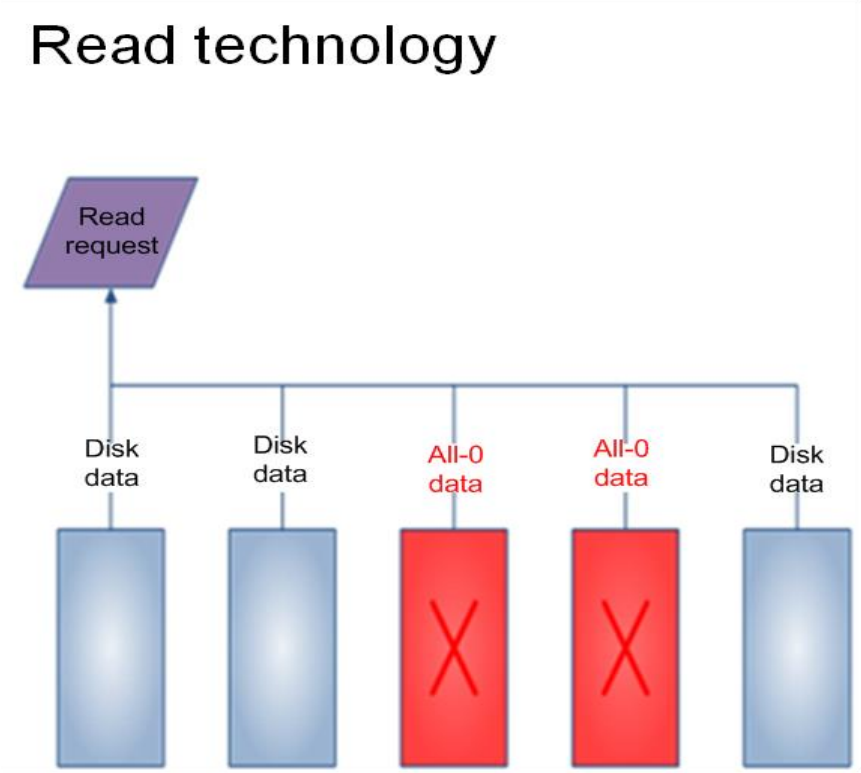
In traditional storage arrays, when a RAID group fails, all data on the RAID group becomes invalid to ensure data consistency. Then the disks cannot be read or written. Huawei designs the SecureVideo technology for video surveillance scenarios. With the technology, reads and writes will not be interrupted during video recording though multiple disks of a RAID group are faulty. Therefore, video data can still be stored and read generally.

As the size of video surveillance systems grows, algorithms in the video surveillance field are also becoming more mature. The reliability and integrity of most video image data can be ensured. Accordingly, in the video surveillance field, maximizing business continuity upon multiple disk failures has become a prevailing demand. SecureVideo is a technology that maximizes business continuity on the basis of sacrificing a small amount of data on faulty disks.

1. SecureVideo read technology

   If the number of faulty disks in a disk domain does not exceed the upper limit, RAID groups can have the same level of reliability as conventional RAID groups. If the number of faulty disks exceeds the upper limit, data on the faulty disks cannot be read but data on functional disks remains readable, ensuring zero service interruption while enhancing the self-repairing possibilities of images of upper-layer services.

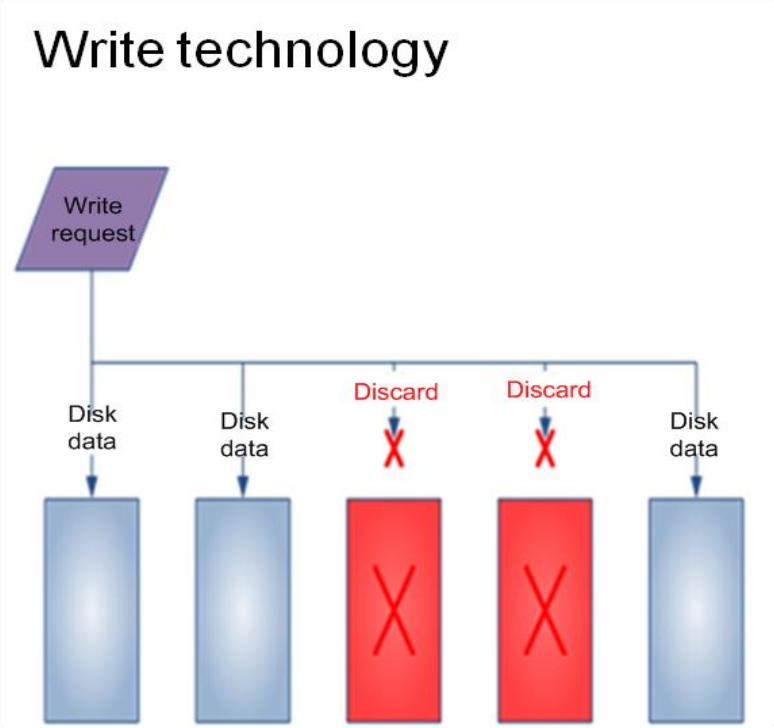
   For example, in a RAID 5 group, if a read request involves five disks and two disks become faulty, data on the faulty disks is filled by 0 and returned as user data along with data on the other three healthy disks. This ensures the request success. For details, see the following figure.



2. SecureVideo write technology

   If the number of faulty disks in a disk domain does not exceed the upper limit, RAID groups can have the same level of reliability as conventional RAID groups. If the number of faulty disks exceeds the upper limit, data written to the faulty disks is discarded while functional disks can still allow data to be written, ensuring zero service interruption and improving service continuity in video recording scenarios. Even if only a proportion of data is written to disks, lost data can still have a chance to be repaired by the image repair algorithm, mitigating the risk of data loss caused by service interruption.

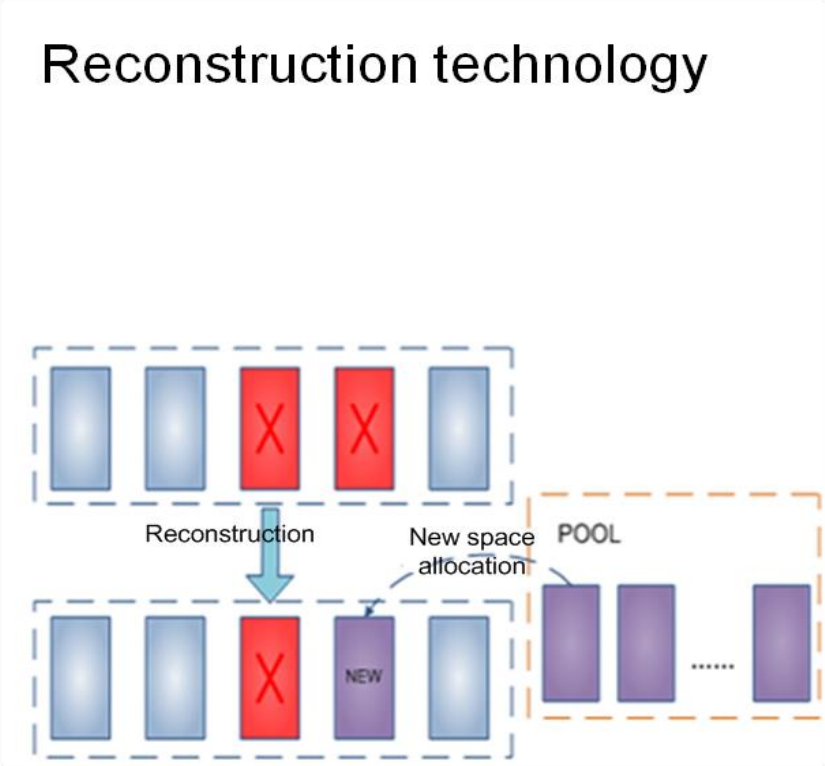
   For example, in a RAID 5 group, if a write request involves five disks but two disks become faulty, data written to the faulty disks is discarded while data can still be written to the other three healthy disks. New user data is largely written and stored to healthy disks and therefore service continuity is achieved. For details, see the following figure.



3. SecureVideo reconstruction technology

   If the number of faulty disks in a disk domain does not exceed the upper limit, data is reconstructed using the conventional RAID redundancy algorithm. If the number of faulty disks exceeds the upper limit, RAID 2.0+ automatically discards data that cannot be reconstructed and replaces the storage space of the faulty disks with that of healthy disks, ensuring the integrity of RAID groups and the reliability of subsequent services.

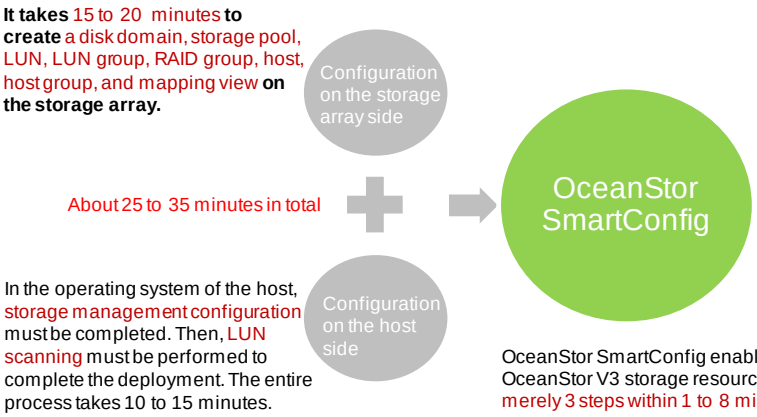
   For example, when two disks of a RAID 5 are faulty, data cannot be reconstructed and is lost. New block space is selected from the storage pool to replace the space of faulty disks to form a new RAID group with healthy disks, ensuring reliability of subsequent services. For details, see the following figure.



3.2.4 SmartConfig Easy Configuration

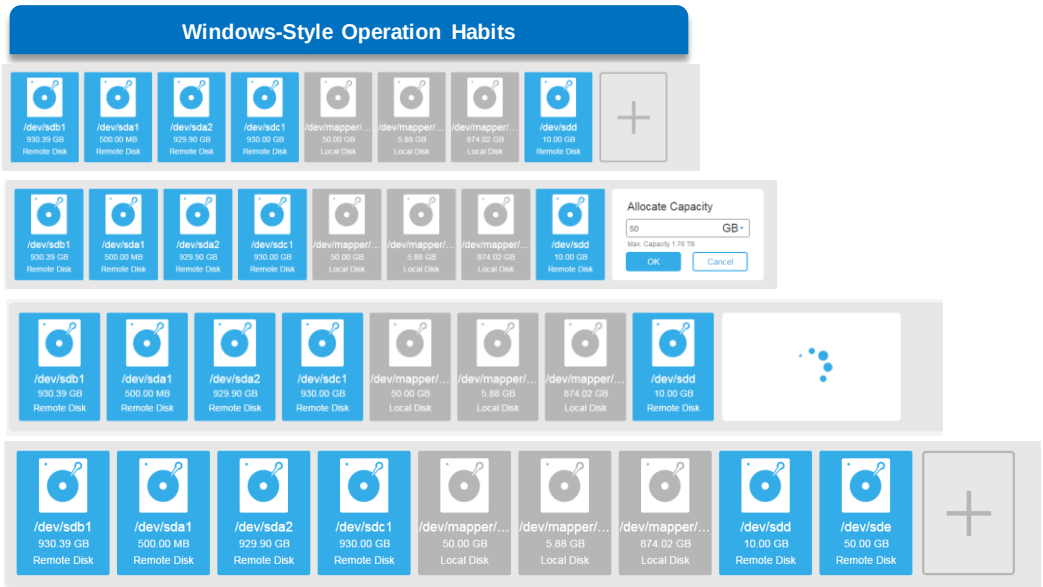
SmartConfig is a piece of software installed on hosts for managing storage devices while simplifying operability and usability of storage devices. This software frees users from complex operations in enterprise-level storage management. You can flexibly and efficiently configure disk resources in storage devices as remote disks and mount them to hosts, and use them like local disks. Traditionally, adding a disk to a host requires complex and sequenced operations on the array through device management software. These operations include creating disk domains, storage pools, LUNs, LUN groups, RAID groups, hosts, host groups, and mapping views. The entire procedure takes 15 to 20 minutes even for someone familiar with storage services. Besides, running commands to log in to the array, scan disks, and perform other operations involves a large number of parameters and therefore is prone to errors. SmartConfig enables resource allocation on the storage system to be completed in three steps, requiring only 1 to 8 minutes.

I. Disks mounted within one minute on one interface



	Traditional Resource Allocation	Resource Allocation Based on SmartConfig
Operation Object	Storage array + Operating system on the host	Tool GUI
Number of Steps	15 steps (11 steps on the storage side + 4 steps on the host side)	3 steps
Deployment Time	25 to 35 minutes	Shorter than 8 minutes

II. Easy-to-use Windows-style operations



3.2.5 Energy-Saving Technologies

I. Intelligent fan speed control

The intelligent fan speed adjustment technology supports 32-level speed adjustment, intelligently tunes the fan speed based on the current system temperature, and reduces fan power consumption and noise, making the system adaptive to various environments.

II. Intelligent CPU frequency adjustment

The intelligent CPU frequency adjustment technology lowers the CPU frequency as the load decreases to minimize the system power consumption.

4 Promotion

4.1 Advantages of Huawei Solution

I. Leading Specifications

- **Leading hardware**: Adopts an active-active dual-controller architecture, new-gen PCI-E 3.0 buses, and 12 Gbit/s SAS 3.0 high-speed disk ports to satisfy video surveillance, checkpoint image capture, and other bandwidth- intensive use cases.
- **Leading performance:** Support for 396 disks on a single device; Supports up to 1024 channels of 2Mbps recording and playback.

II. Safe and Reliable

- **Online scalability**: Controllers, fans, power supplies, disks, and interface modules are all hot-swappable; resources are linearly expandable online; and disk roaming technology decouples disks from slots to ensure services remain up and running.
- **Data protection**: Innovative RAID2.0+ data protection technology achieves automatic load balancing among hard disks through virtualization of the underlying layer, reducing the disk failure rate and delivering a 20-fold improvement in data reconstruction speeds compared to traditional models.

III. Ease of Use

- **Easy configurations**: Huawei proprietary SmartConfig automatically completes server mounting, LUN configurations, and other operations; storage resource configurations completed in three simple steps in less than one minute.
- **Easy to manage**: GUI-based management tools; visualization of capacity, performance, fault, and other information; and support for a wide range of alarm handling procedures helps simplify management.

IV. Energy-Saving

- **Intelligent 32-speed fan controls**: Fan speeds are intelligently adjusted according to system temperature, reducing power consumption and noise while saving on operating costs.
- **Efficient power supply**: Uses a platinum power supply with up to 94% power conversion efficiency.

5 Conclusion

Huawei OceanStor 2600 V3 for Video is dedicated to the video surveillance field and characterized by cost-effectiveness, excellent performance, large capacity, high reliability, flexible scalability, and easy maintenance. It can help customers lower the TCO of video surveillance applications, simplify management, and improve performance.

A Acronyms and Abbreviations

Acronyms and abbreviations

Acronym and Abbreviation	Full Spelling
CCTV	Closed Circuit Television
DVR	Digital Video Recorder
IPVS	IP Video Surveillance
VCR	Video Cassette Recorder
IP	Internet Protocol
FC	Fibre Channel
JPEG	Joint Photographic Experts Group
MPEG	Moving Picture Experts Group
SAN	Storage Area Network
RAID	Redundant Array of Independent Disks
LUN	Logical Unit Number
SATA	Serial Advanced Technology Attachment
ISM	Integrated Storage Management
TCO	Total Cost of Ownership