Huawei OceanStor 2800 V5 Storage System (Video Surveillance Edition) Technical White Paper

Issue 01

Date 2017-11-14

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions

Notice

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

Contents

1 Abstract

As social economy develops rapidly, a series of problems involving social security, traffic, law enforcement in cities are arising. Under this situation, 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, number of high definition (HD) cameras is rapidly increasing. In addition, image quality has been developed from Common Intermediate Format (CIF) to 720P and 1080P. 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.

Huawei products can help customers meet needs of various video surveillance applications at a lower price and with higher performance.

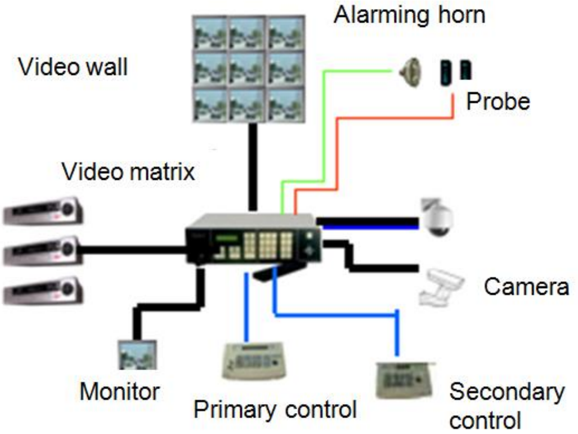
2 Overview

2.1 Industry Trend

2.1.1 1st-Generation Video Surveillance System

Video surveillance goes through three phases: 1st-generation analog video surveillance system (CCTV), 2nd-generation digital video recorder (DVR) system, and 3rd-generation network video surveillance system (IPVS). With the development of 3rd-generation video surveillance system, open direct storage of video streams gradually appears, simplifying the system architecture and reducing the total cost of ownership (TCO).

Figure 2-1 1st-generation analog video surveillance system



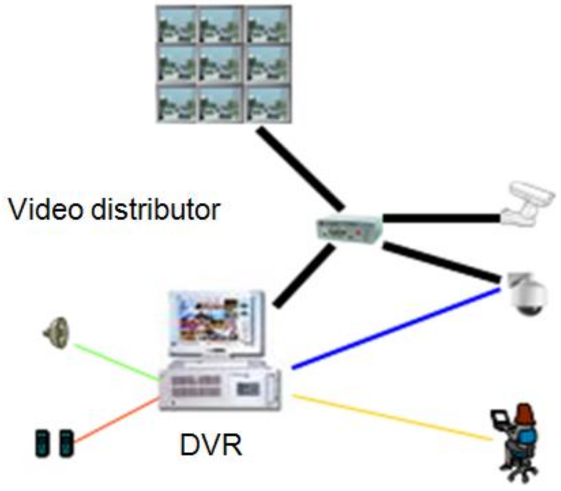
The 1st-generation analog video surveillance system requires dedicated devices including cameras and monitors. For example, by sending a video signal, a camera connects to a dedicated analog video device such as a video multiplexer, video matrix, video cassette recorder (VCR), or video monitor. Many limitations exist in the 1st-generation analog video surveillance system. For example:

- Surveillance capabilities are limited. Only local video surveillance is supported.
- Storage capabilities are restricted by the video matrix capacity.
- Users whose video loads are heavy must replace videotapes for cameras in a timely manner. The quality of videos that are saved in videotapes cannot meet requirements.

The quality decreases as the number of copies increases. Videotapes are difficult to save. They occupy large storage capacities and are prone to be lost or unintentionally erased.

2.1.2 2nd-Generation Video Surveillance System

Figure 2-2 2nd-generation DVR system

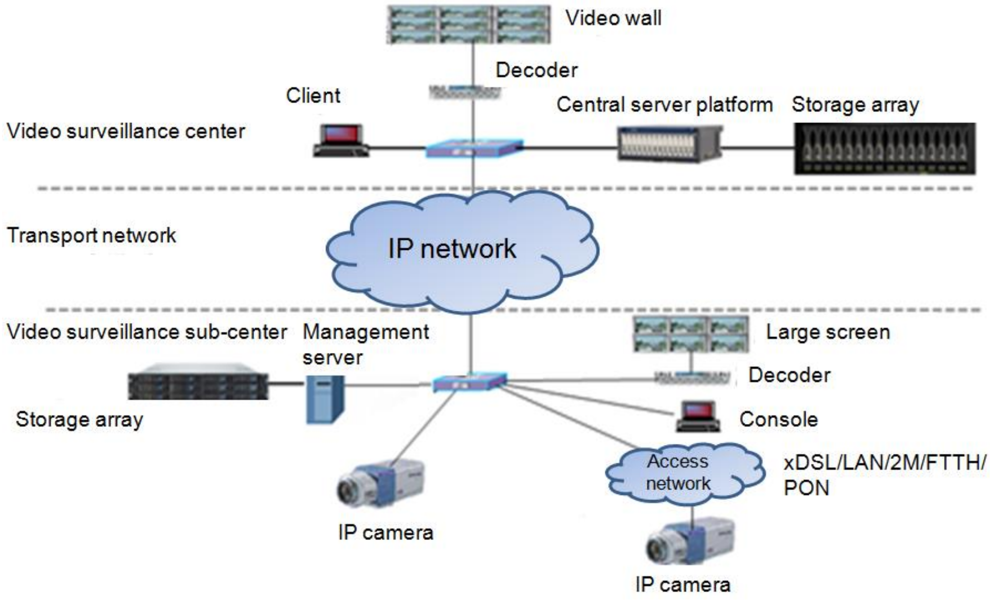


The 2nd-generation DVR system is a semi-analog and -digital video surveillance system that uses DVRs as core devices. Cameras and DVRs employ coaxial cables to transfer video signals. The 2nd-generation video surveillance system supports video recording, playback, and limited IP network accesses based on DVRs. However, a variety of DVRs exist and they do not have a unified standard. Therefore, the 2nd-generation DVR system is a non-standard closed system. Many limitations exist in the system. For example:

- Cabling is complex. Independent video cables must be connected to each camera.
- Scalability is restricted. A maximum of 16 cameras can be expanded at a time.
- Manageability is restricted. An external server or management software is required to manage multiple DVRs.
- Cameras do not support direct remote access. DVRs are required to access the cameras.
- Disk redundancy protection is unavailable. Data is easy to lose when a disk fails.

2.1.3 3rd-Generation Video Surveillance System

Figure 2-3 3rd-generation network video surveillance system



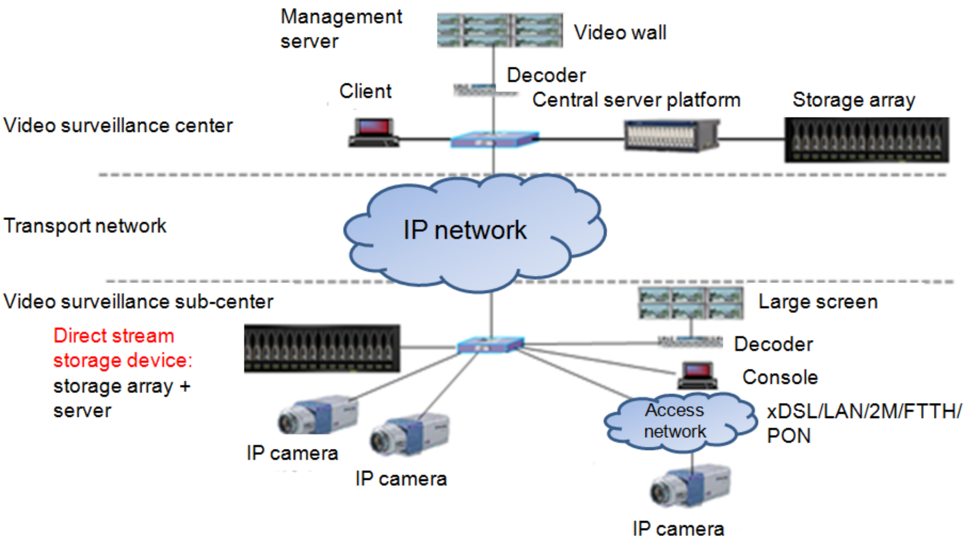
The 3rd-generation video surveillance system is an all-IP video surveillance system. Being compared with the first two generations, it features built-in web servers of cameras and provides Ethernet ports for access. The JPEG or MPEG4 data files that are generated by cameras can be accessed, monitored, and replicated by all authorized clients at all positions on a network. The 3rd-generation video surveillance system outclasses the first two generations in the following aspects:

- Easy connection: Cameras can be connected to the live network by setting the local area network (LAN).
- Centralized control: The entire video surveillance can be managed by a standard server and a set of management software.
- Flexible expansion: Cameras can be flexibly added and servers can be upgraded easily. Faster processing and larger disk capacity are available.
- Remote management: All authorized users can directly access and manage cameras over a network as well as user a server to centrally manage cameras.
- Professional storage: The storage systems support IP SAN and FC SAN, meeting requirements for larger capacity and faster processing. The RAID technology ensures data reliability.

2.1.4 4th-Generation Video Surveillance System (Based on Open Direct Storage of Video Streams)

The 3rd-generation video surveillance system is implemented in central server + storage server mode. The reliability of the central server is inferior to that of professional storage devices. The central server is a source that may cause single points of failure. In addition, physical servers occupy more space and consume more power. Cabling is complex. After the 3rd-generation video surveillance system is optimized, the video surveillance system based on open direct storage of video streams comes into being. In this system, the central server and storage server are integrated into one direct stream storage device. The direct stream storage device can directly write the received video surveillance data onto disks instead of using an independent central server to transfer the data.

Figure 2-4 Video surveillance system based on open direct storage of video streams



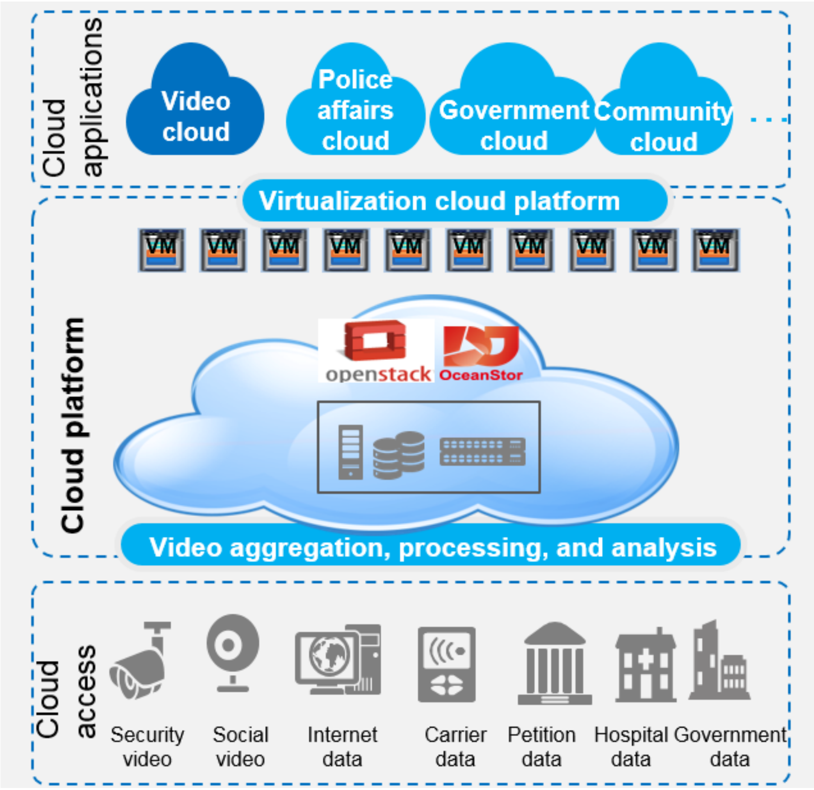
Huawei uses the mainstream virtual machine (VM) technology to combine the central server and storage array into one device, that is, integrated direct stream storage device. The central server runs on a controller of the storage array in the form of a VM. An internal high-speed memory channel is used between the server and storage array. The video surveillance system based on open direct storage of video streams is superior to the first three generations in the following aspects:

- Easy connection: Physical cables and network adapters are no longer needed to connect a server to a storage array.
- Centralized control: The direct stream storage device can be used to directly manage storage arrays and the video surveillance server.
- Flexible expansion: Retaining robust reliability, outstanding performance, and flexible scalability of traditional storage arrays, the direct stream storage device can integrate more third-party video surveillance applications using VMs.
- Energy saving: Servers exist in a virtual mode, minimizing the power consumption of the entire solution.
- Small space occupation: Servers do not need physical storage space.
- Cost effectiveness: Storage arrays give full play to their hardware capabilities, reducing the TCO.
- Optimal performance: Front-end 10GE network adapters and back-end 12 Gbit/s SAS ports are used for interconnection and the latest PCIe 3.0 technology is applied to the mirror channel, making better user experience available.

2.1.5 5th-Generation Video Surveillance System

With the convergent, efficient, reliable, and data sharing features of the cloud system, the 5th-generation video surveillance system introduces cloud concepts into the video surveillance field. Based on technologies such as computing cloud, storage pooling, and unified management and control, the video surveillance system provides professional information infrastructure, which is open and compatible with more public security services.

Figure 2-5 Video surveillance system based on cloud



Compared with traditional video surveillance systems, the video surveillance cloud has three advantages: full sharing, fully open, and unified management.

- An efficient share-it-all platform

  The "One Cloud One Pool" solution provides a basic platform to facilitate consolidation of upper-layer businesses and data, enable elastic scalability of IT resources, and eliminate the need for data migration.

  Cloud-based IT resource management eliminates the silo construction mode, improving upper-layer application efficiency by 30%. Video image and analysis resources can be obtained on demand across departments and regions to support practical operations and investigation of major cases.

- Open platform

  Software is decoupled from hardware and data is decoupled from applications to build an open and standard platform. Users can choose different vendors of platform components to obtain the optimal solution portfolio.

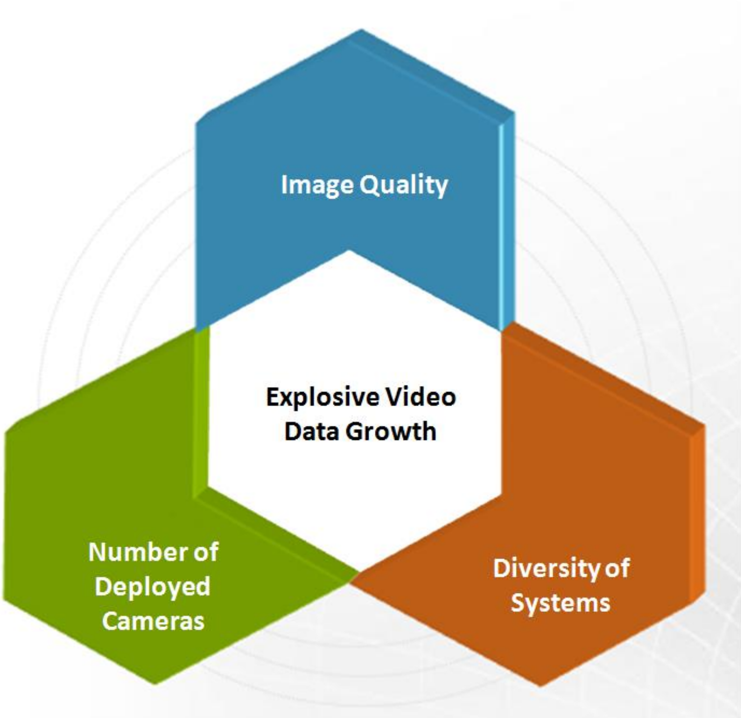
  The open platform will meet the requirements on system inheritance, scalability, and future innovation.

- Unified management and O&M

  Visualized unified management of computing, storage, and network devices and resources enables cross-cloud, cross-department and cross-region management.

2.2 Development Trends of Video Surveillance Systems

2.2.1 Explosive Data Growth



- Image quality of surveillance

  Number of high definition (HD) cameras is increasing. The resolution of cameras evolves from low definition (LD) (2 Mbit/s) to HD 720P (4 Mbit/s) and full HD 1080P (6 or 8 Mbit/s). The storage capacity is increased by three to four times. Currently, HD and full HD cameras dominate the video surveillance industry.

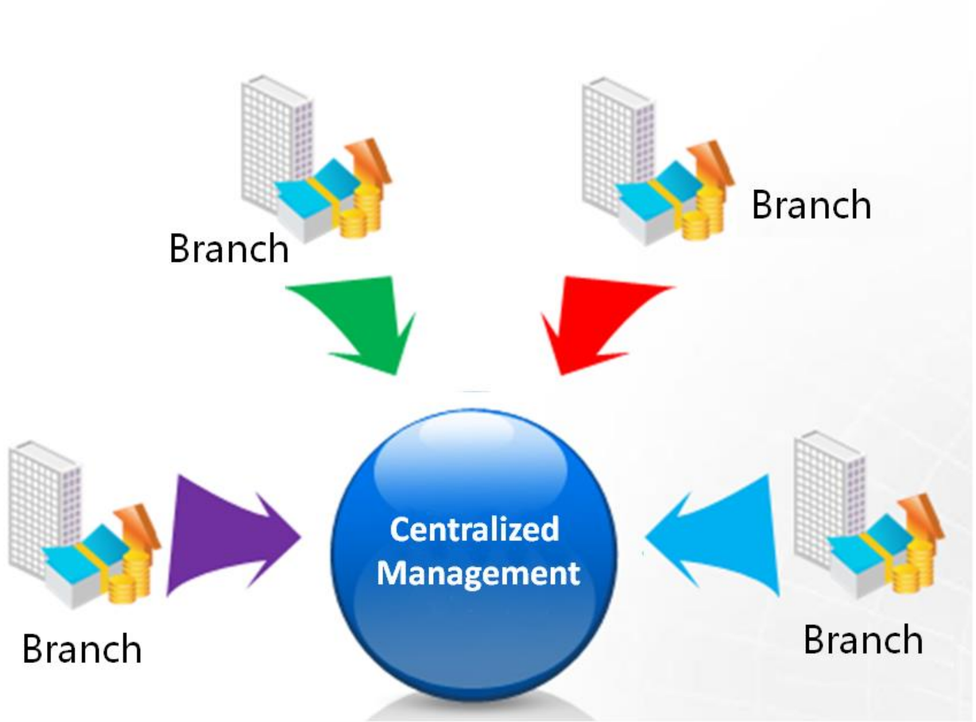
- Camera sampling points

  Video surveillance systems are widely used in fields such as urban traffic, Safe City projects, and unattended areas. Camera sampling points and capacities are also increasing exponentially.

- Diversity of systems

  Video surveillance systems are interconnected with access control systems, alarming systems, and other systems. Data storage capacities are increasing to meet file backup and archiving needs.

2.2.2 Popularity of Centralized Management



As video camera sampling points and video 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.

2.2.3 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, video surveillance's demand for large-capacity storage systems is increasing. To satisfy this demand, major storage vendors have all released large-capacity storage systems above PB level.

2.2.4 Full Interconnection and Intelligence Requirements

Construction of an all-round urban public security protection and control system requires a 24/7, comprehensive, full-HD, intelligent, and multi-dimensional Smart City center to support security control, urban management, and civil services, and to effectively improve the efficiency of public security services. Video surveillance is actively working with cutting-edge technologies such as cloud computing, big data, and intelligent video analysis to build a cloud service management system and a data resource sharing pool. Based on these, a service platform that provides view management, view resolution, and video big data capabilities can be established, forming a new video sharing mode to be applied in practice.

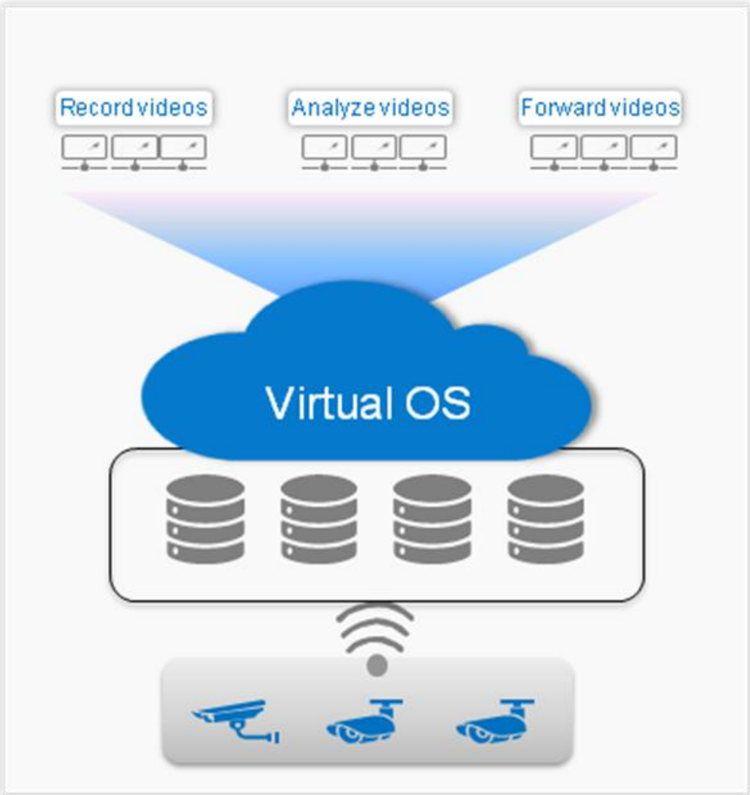
3 OceanStor 2800 V5

3.1 Introduction to OceanStor 2800 V5

3.1.1 Overview

Huawei OceanStor 2800 V5 is the next-generation video surveillance storage system dedicated to access monitoring centers. The OceanStor 2800 V5 is an all-in-one device that integrates video surveillance services. Based on the next-generation V5 storage arrays released by Huawei, the OceanStor 2800 V5 has built-in VMs and uses VMs to carry the excessive computing capabilities of storage controllers to migrate third-party applications that were deployed on physical servers to the controllers in the storage array. Figure 3-1 shows the logical architecture of the OceanStor 2800 V5.

Figure 3-1 Logical architecture of the OceanStor 2800 V5



While retaining IP SAN access capabilities, the OceanStor 2800 V5 has excellent scalability, reliability, and performance. VMs are used to migrate server capabilities to the storage controllers, helping customers lower their total cost of ownership (TCO) and save more equipment room space.

Compared with Huawei OceanStor S2600T, the last generation of storage systems dedicated to video surveillance, the OceanStor 2800 V5 employs the RAID 2.0+ technology, which is solely developed by Huawei. The RAID 2.0+ technology can effectively reduce the risks of dual-disk failures and brings higher performance to the storage system. In addition, workloads can be automatically evenly allocated to all the disks in a disk domain, and management of the storage system is easier with no specified hot spare disks required.

3.1.2 High-Density Disk Enclosure

High-density disk enclosures fully utilize their 4 U space. Vertical instead of horizontal disk slots are provided, and the size of cascading boards is reduced by more than half, providing as much space as possible for disks. The delicate ventilation channel design, optimized fan speed adjustment policy, and enhanced fan usage resolve the heat dissipation issue and meet the noise and power consumption requirements. Currently, one disk enclosure can house up to seventy-five 3.5" disks.

Compared with common 4 U 3.5" disk enclosures, high-density disk enclosures have the following advantages:

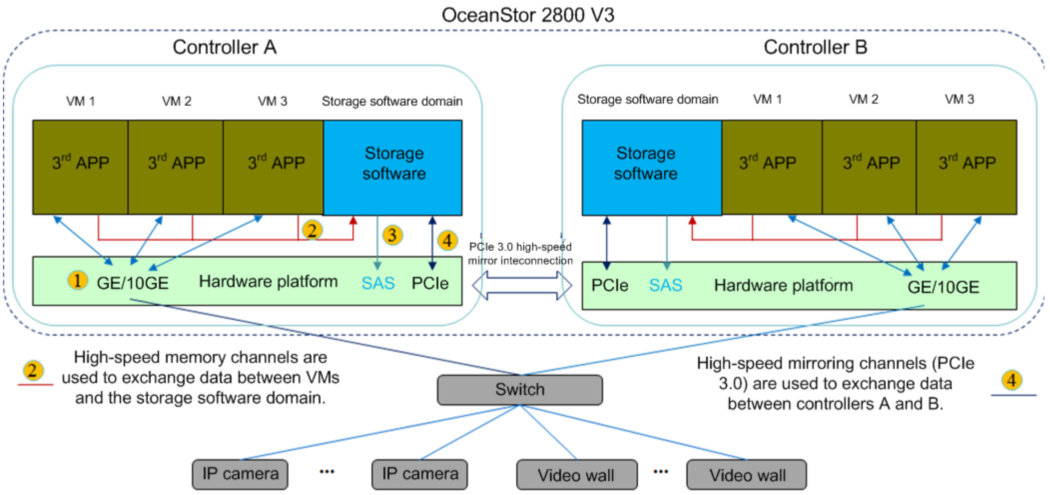
- The disk density of high-density enclosures is three times that of traditional disk enclosures.

- The performance of each U in high-density disk enclosures is twice that of traditional disk enclosures.
- The operating expense per disk of high-density disk enclosures is 20% lower than that of traditional disk enclosures.
- The power consumption of each disk in high-density disk enclosures is 50% lower than that of traditional disk enclosures.

3.1.3 Built-in Storage VMs (SmartContainer)

Through virtualization technologies, storage controllers configure storage software as an independent domain and run third-party applications on VMs. This not only brings excellent isolation but also gives full play to the computing capabilities of storage systems.

Figure 3-2 Architecture of built-in storage VMs



1. Applications receive surveillance video data. Third-party applications (video surveillance applications) deployed on VMs use GE or 10GE IP networks to receive video data from IP cameras. Services use their dedicated ports to avoid service interference. Each controller supports up to 12 GE network ports or eight 10GE network ports.
2. Surveillance video data is written into storage. Data is exchanged between VMs and storage software domains through internal high-speed memory channels. DDR3-1600 memory modules are used. Compared with the physical IP networks originally used between servers and storage systems, memory channels bring higher performance and do not require cabling.
3. Surveillance video data is written onto disks. Compared with the 6 Gbit/s SAS connection used in the OceanStor 2600T, the last generation of storage systems, the OceanStor 2800 V5 adopts next-generation 12 Gbit/s SAS interconnection, which improves performance and reduces latency for video data read and write.
4. Surveillance video data is mirrored. To ensure high data reliability, the OceanStor 2800 V5 uses dual controllers. Mirroring channels between controllers A and B use the latest PCIe technologies, with the two-way bandwidth reaching as high as 8 GB/s.

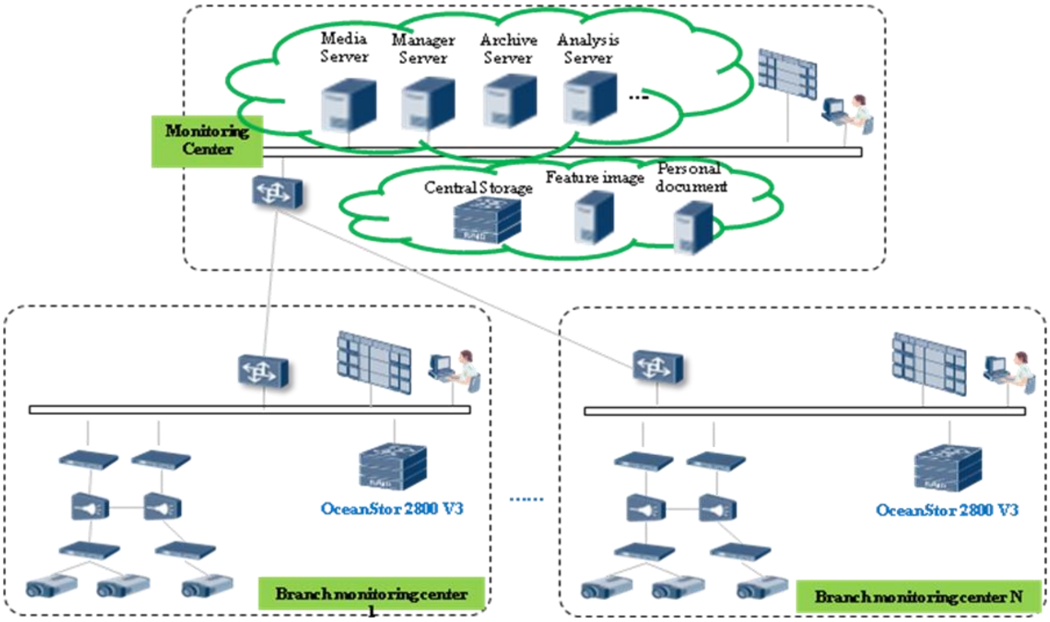
With excellent compatibility and isolation brought by virtualization technologies, storage services, forwarding services, index services, and other third-party applications of video surveillance can be deployed on storage controllers to reduce TCO.

3.1.4 Typical Networking

As shown in Figure 3-3, the typical network of the OceanStor 2800 V5 adopts centralized storage.

Centralized storage systems are well suited to small- and medium-sized networked video surveillance projects. In these projects, only one or two OceanStor 2800 V5 systems are required, saving much equipment room space. Apart from saving video surveillance data, the OceanStor 2800 V5 can also be accessed through IP SAN to provide external storage services. Therefore, it can minimize customers' investment.

Figure 3-3 Typical network of the OceanStor 2800 V5 video surveillance storage solution



Solution highlights:

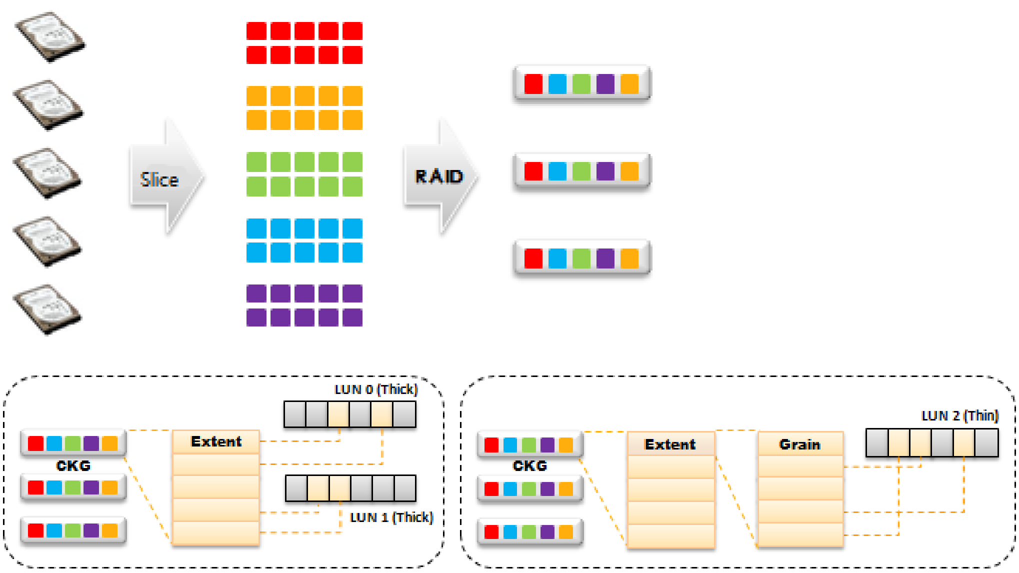
- Suits scenarios where centralized storage/monitoring is required with a small number of users, for example, areas or tier-2 storage nodes on large networks.
- Minimizes customers' initial investment, and follow-up expansion and maintenance costs.
- Eliminates the necessity of multiple physical servers and consumes less space and power.
- Enables excessive storage capabilities to be accessed through IP SAN to provide external storage services.
- Enables customers to integrate or separate service and management networks based on service requirements.

3.2 Huawei Video Surveillance Storage Solution

3.2.1 RAID 2.0+

Huawei OceanStor 2800 V5 is the key component in Huawei video cloud access solution, and industry-leading RAID 2.0+ technology is available on the OceanStor 2800 V5. RAID 2.0+ adopts a 2-layer virtual management model that consists of underlying disk management and upper-layer resource management. In the system, all disk space is divided into data blocks with a smaller granularity. Based on the data blocks, RAID groups are built to evenly distribute all data on all disks in the storage pool. With the disk virtualization redundancy protection technology, technical restrictions of traditional RAID groups are eliminated. Partial data reconstruction, thin reconstruction, and parallel reconstruction can be performed, with the data reconstruction speed reaching as high as 2 TB/hour. Resources are managed based on data blocks, resulting in higher resource management efficiency.

Figure 3-4 Working principles of RAID 2.0+



1. OceanStor storage systems support SSDs, SAS disks, and NL-SAS disks, and these disks are built into disk domains. In a disk domain, disks of the same type compose disk groups (DGs) based on certain rules.
2. The space of each disk in a DG is divided into fix-sized chunks (CKs). Using random algorithms, OceanStor storage systems create chunk groups (CKGs) out of the chunks from different disks based on RAID algorithms.
3. Each CKG is divided into fix-sized logical storage spaces called extents. Extent is the basic unit of thick LUNs (also called fat LUNs). To create thin LUNs, extents are further divided into grains that have a smaller granularity and then mapped to thin LUNs with the unit of grain.

For more details about Huawei RAID 2.0+ technology, see the *Huawei OceanStor Enterprise Class Storage Systems RAID 2.0+ Technical White Paper V1.2*.

3.2.2 Reliability Technologies

3.2.2.1 System Reliability: Active-Active Dual-Controller Architecture

The OceanStor 2800 V5 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.

3.2.2.2 System Reliability: Cache Mirroring

The OceanStor 2800V5 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.

3.2.2.3 System Reliability: RAID 6

As NL-SAS and SATA disk capacities increase, 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 5 disks, the probability will fall to 0.000,026%, 1,995 times less. 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).

3.2.2.4 Disk Reliability: Disk Pre-reconstruction

Disk pre-reconstruction refers to that a 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-reconstruction technology eliminates the risks of both RAID degradation and data loss.

3.2.2.5 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.

3.2.2.6 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. Further, the technology reduces RAID group failures due to bad sectors occurring on multiple disks in a disk pool all at once.

3.2.2.7 Disk Reliability: Online Diagnosis

The online diagnosis technology enables failed disks to be recovered without interrupting ongoing services. When faults or read or write errors occur on a disk 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 on and off 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 causes the RAID group to fail:

   Temporarily suspend data reads and writes on the failed disk, and retry it or power it on and off to recover it.

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.

3.2.2.8 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

3.2.3.1 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 OceanStor 2800 V5 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 OceanStor 2800 V5 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 OceanStor 2800 V5 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 is automatically recovered using RAID algorithms, and the reconstructed data is rewritten into a selected hot spare space.

3.2.3.2 Video Image Repair (SecureVideo)

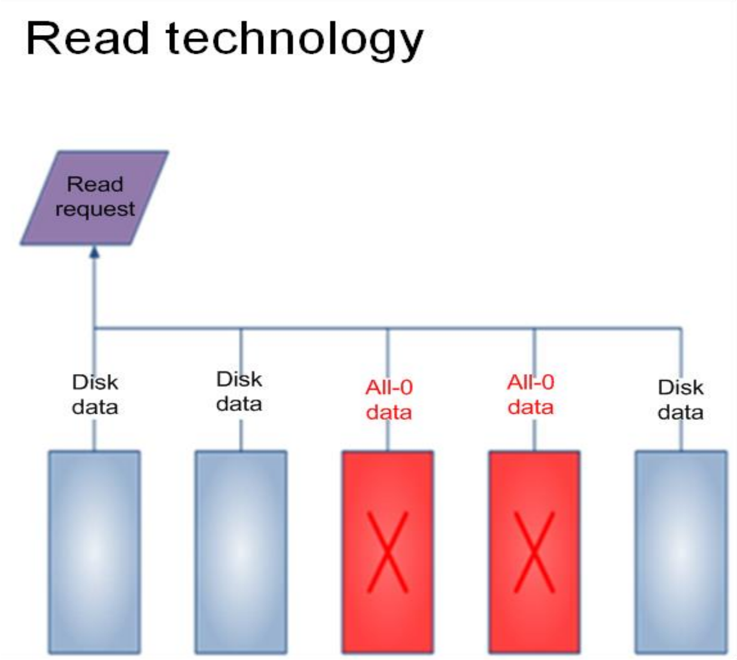
In traditional storage arrays, when a RAID group fails, all data on the RAID group will fail to ensure that the data is consistent. However, Huawei has designed an image repair technology catered for video surveillance scenarios. With this technology, if a RAID group fails, data cannot be written to the bad disks, but data on the remaining working disks can still be read. Therefore, video data on the disks is still usable. If multiple disks of a RAID group fail, neither writes nor reads of video surveillance data are stopped.

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. In contrast, 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.

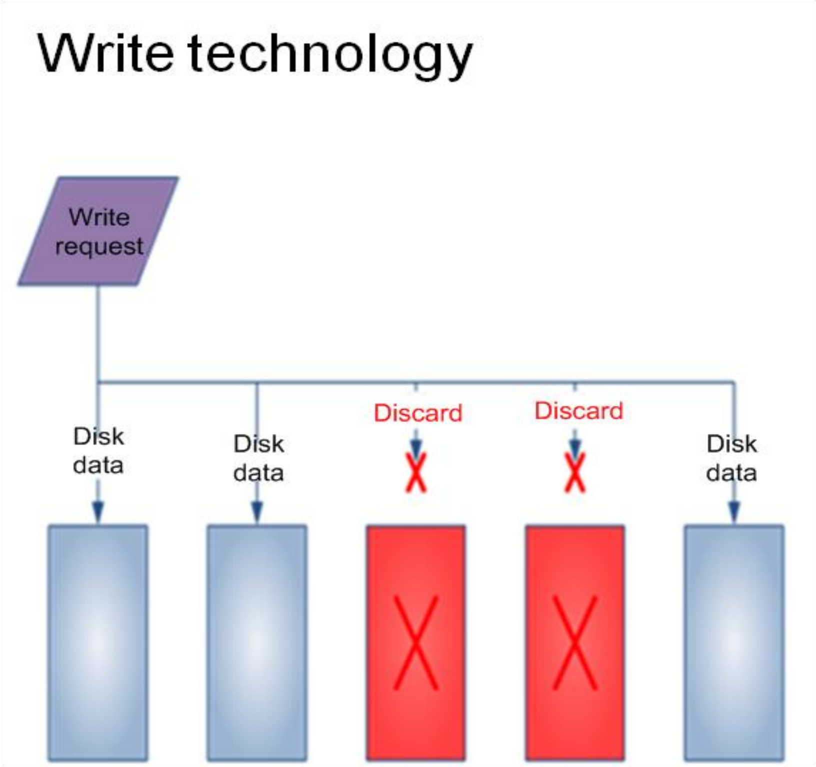
   For example, in a RAID 5 group, if a read request involves five disks and two disks become faulty, data on the faulty disks is filled by 0 and returned as user data along with data on the other three healthy disks. See the following figure:



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. In other words, 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 risks of data loss caused by service interruptions.

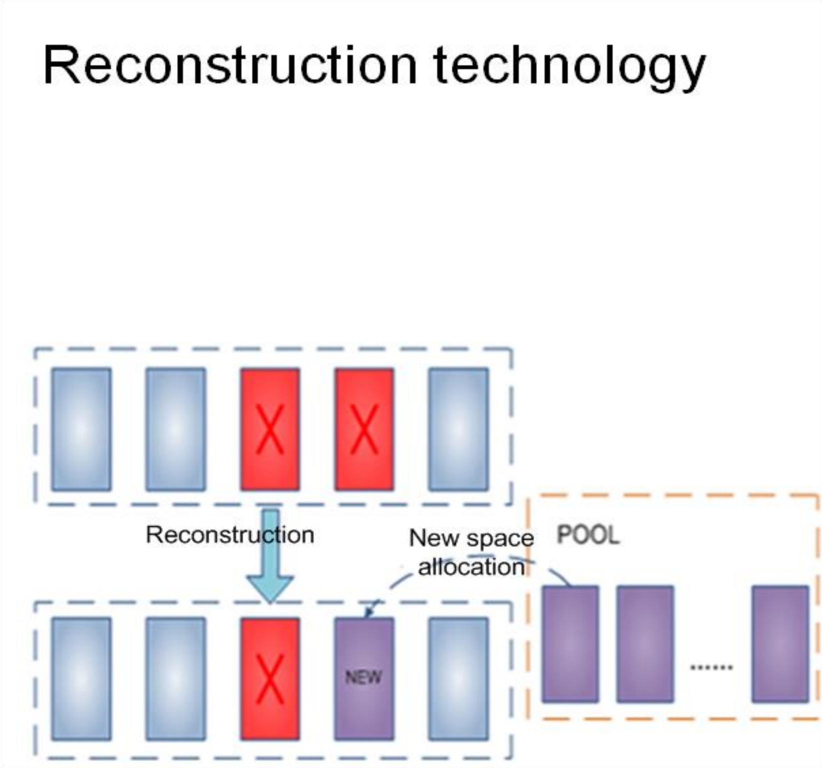
For example, in a RAID 5 group, if a write request involves five disks but two disks become faulty, data written to the faulty disks is discarded while data can still be written to the other three healthy disks. New user data is largely written and stored to healthy disks and therefore service continuity is achieved. 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, in a RAID 5 group, if two disks become faulty, their data is lost when the data cannot be reconstructed based on RAID properties. To ensure the reliability of subsequent services, the space of new blocks from the storage pool is used to replace that of the faulty disks and form a new RAID group together with the space of the other healthy disks. See the following figure:



NOTE

This technology is only available for SAN services on OceanStor 2800 V5 storage systems.

3.2.4 Energy-Saving Technologies

3.2.4.1 Intelligent Fan Speed Adjustment

The intelligent fan speed adjustment technology supports 16-step speed adjustment, intelligently tunes the fan speed based on the current system temperature, and reduces fan power consumption (which accounts for 15% of overall system power consumption) and noise, making the system adaptive to various environments.

3.2.4.2 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

- Large capacity

  An OceanStor 2800 V5 storage system can support 750 disks.

- Outstanding performance

  The OceanStor 2800 V5 inherits the high performance, high reliability, and excellent scalability from the OceanStor V5 series. A single set of the OceanStor 2800 V5 (with dual controllers) can support a maximum of 1600 channels of 4 Mbit/s video recording and 400 channels of 4 Mbit/s video playback, which is an unparalleled performance advantage over products from other video surveillance storage vendors.

- Cost-effectiveness

  The OceanStor 2800 V5, allowing third-party applications to be integrated using VMs, is specially designed and optimized for the video surveillance field. With outstanding performance at a lower price, the OceanStor 2800 V5 helps customers tremendously cut their TCOs.

- Flexible scalability

  The OceanStor 2800 V5 supports flexible networks, including the video surveillance storage networks and IP SAN networks. The OceanStor 2800 V5 has a variety of front-end host ports. By default, every controller is configured with four Gigabit Ethernet (GE) ports and interface modules can be added based on customers' needs. A maximum of 10 high-density disk enclosures can be configured at the back-end, supporting easy online disk expansion.

- Superb reliability

  The OceanStor 2800 V5 uses Huawei proprietary RAID 2.0+ technology and many other reliability technologies to prevent data losses upon a system failure and fully protect customers' assets.

- Easy maintenance

  The OceanStor 2800 V5 uses a uniform command-line interface (CLI) interface. The configuration and maintenance of the OceanStor 2800 V5 and the life cycle management and configuration of VMs can all be done in the CLI. The OceanStor 2800 V5 features easy configuration, maintenance and one-click upgrade.

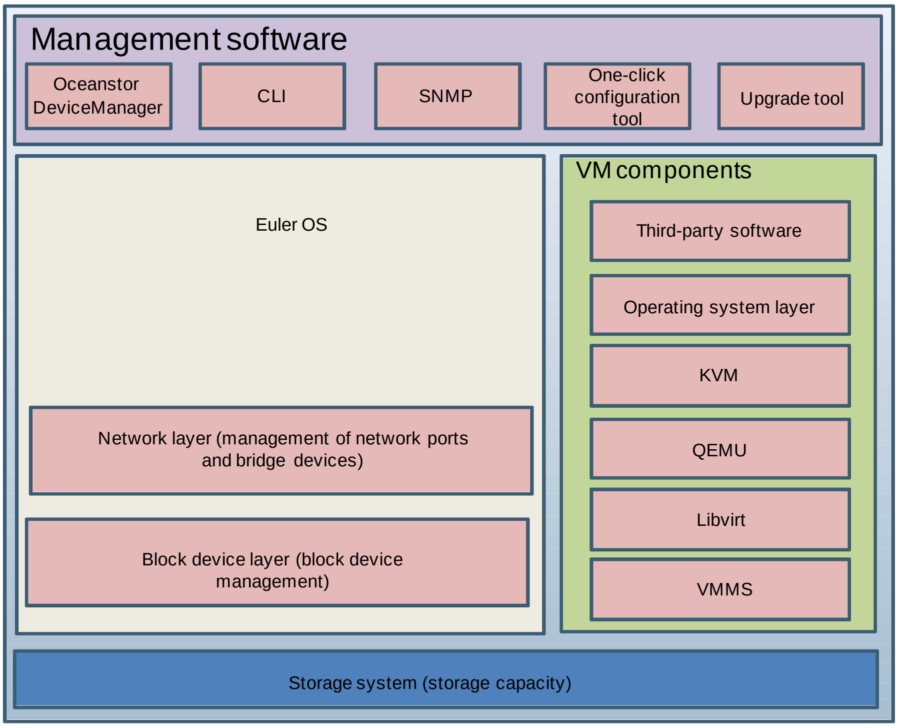
5 SmartContainer Feature

5.1 SmartContainer Introduction

It is a new attempt to embed virtual machine (VM) technology in storage systems. Different from hosts with VMs embedded, systems converging storage and virtualization can use their own storage resources, thereby reducing maintenance costs and simplifying service configuration. The mainstream VM components are VMware, Xen, KVM, and Docker.

5.2 Basic Principle

SmartContainer, as a subsystem in a storage system, provides stable and reliable operating environments for applications on upper-layer hosts based on Huawei unified storage. KVMs are embedded in the storage system and applications are deployed on operating systems running on the VMs to achieve the preceding purposes. The Libvirt component is used to manage the VMs.



Each controller supports a maximum of three VMs. The VMs run on spare CPUs and memory resources, use physical network ports of storage systems as network resources, and employ LUNs in storage systems as storage resources.

The vHost front-end driver is added to the TGT module to convert LUNs in a storage system into block devices. The block device mapping mechanism provided by the KVM maps block devices on hosts to VMs as data disks. The EAM module reports maintenance status and manages mutual exclusion of ports.

5.3 Major Functions of SmartContainer

5.3.1 VM Lifecycle Management and Flexible Resource Configuration

SmartContainer employs the Libvirt component to configure and manage VMs and allows you to:

Create, delete, modify, power on, power off (forcibly power off), restart (forcibly restart), and suspend VMs.

Flexibly adjust and configure resources for VMs.

Allocate and reclaim some resources (network ports and data disks) even when VMs are running.

5.3.2 High Availability of VMs

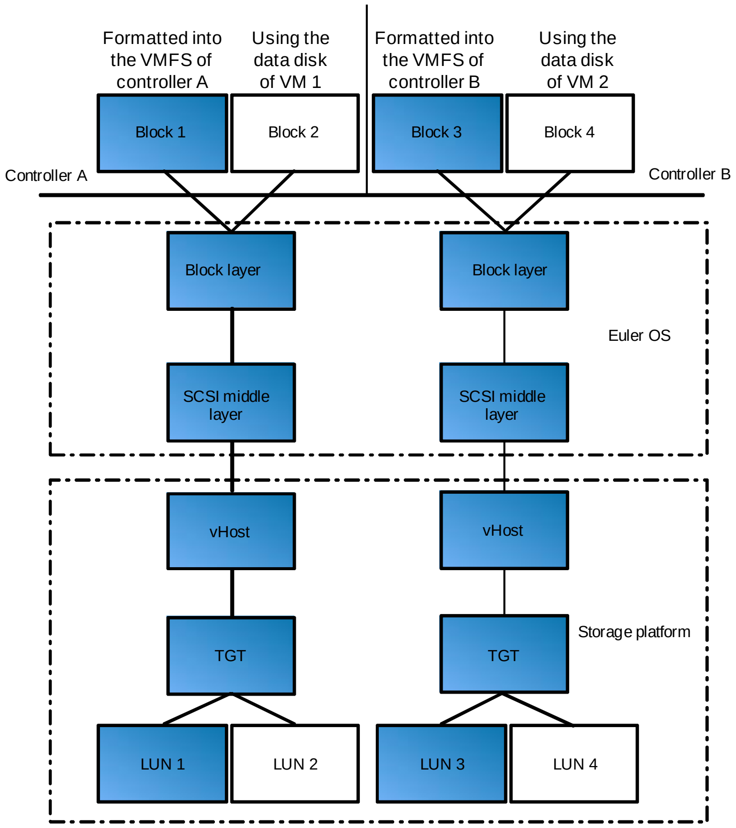
To minimize service downtime, SmartContainer enables you to configure VM failover. After the failover function is enabled, if the active controller is faulty, the VMs fail over to the standby controller for service continuity and fail back to the active controller after the fault is rectified.

5.3.3 VM File System Management

A virtual machine file system (VMFS) is an ext3 file system generated by formatting a specified block device on a host. A block device is converted by vHost from a created LUN. A VMFS is used to store large system disk files on each VM in the node and VM template image files, operating system ISO image files, and VMTools drivers imported or temporarily generated. SmartContainer provides a simple file management mechanism to help you manage these files with ease.

OceanStor 2800 V5 uses the first four (or eight) SAS disks to create the VM file system.

The following figure illustrates the relationship between file systems and LUNs.



5.3.4 Resource Isolation

The CPU and memory resource isolation function is provided to reduce the impact of VMs on storage services. It enables you to configure the number of CPU cores used by VMs and storage services as well as the CPU isolation switch. When the CPU isolation switch is enabled, VMs and storage resources can only run on CPU cores allocated to them and in different memory spaces. In the mean time, VM properties ensure that VM resources are isolated from each other.

Device Model	Memory Specifications
OceanStor 2800 V5	By default, the VM function is enabled. A single controller provides 48 GB memory and 16 logical cores. VM memory usage: 19 GB Number of CPUs available for VMs: 1 to 8 (configurable)

5.3.5 Batch VM Creation Using a Template

The system installation and service deployment processes are complex each time after VMs are created. SmartContainer enables the system disk of an installed VM to be exported and stored as a template and allows VMs to be created using the template, remarkably reducing the time required for batch or cross-controller deployment of services.

5.4 Highlights of SmartContainer

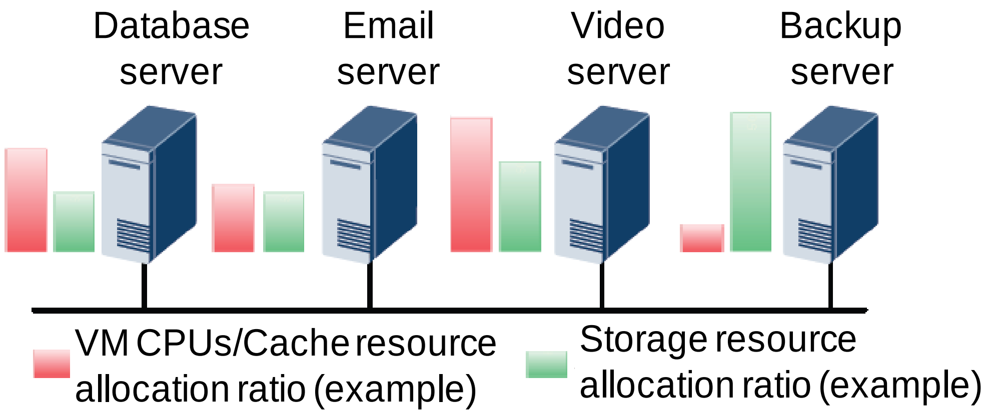
5.4.1 Isolation of Storage from Virtual Resources

To give device performance into full play, you can configure CPU cores for VMs and storage services or even directly disable the CPU isolation switch after configuring VMs so that the storage services and VMs share all CPU cores, helping you find the optimal configuration in varying workload scenarios.

5.4.2 Flexible Allocation of Virtual Resources

The virtual machine (VM) technology becomes popular in many scenarios because it greatly improves application server utilization and lowers the deployment cost and operation expenditure (OPEX) of services. However, VMs are facing a high-density problem as an increasing number of applications and virtual desktops are deployed in VMs. A higher VM density indicates several-fold increase in service data volumes and consumed data bandwidth than a single server. This poses ever demanding requirements on storage capacity, performance, and scalability.

It is a tendency nowadays for one storage system to process diversified applications. However, those applications have differentiated storage requirements, so the storage system must have high flexibility in resource allocation. SmartContainer allocates different resources (memory resources and CPU cores) to VMs running different applications and allows you to flexibly adjust CPU cores consumed by VMs in each controller to attain the optimal device performance.



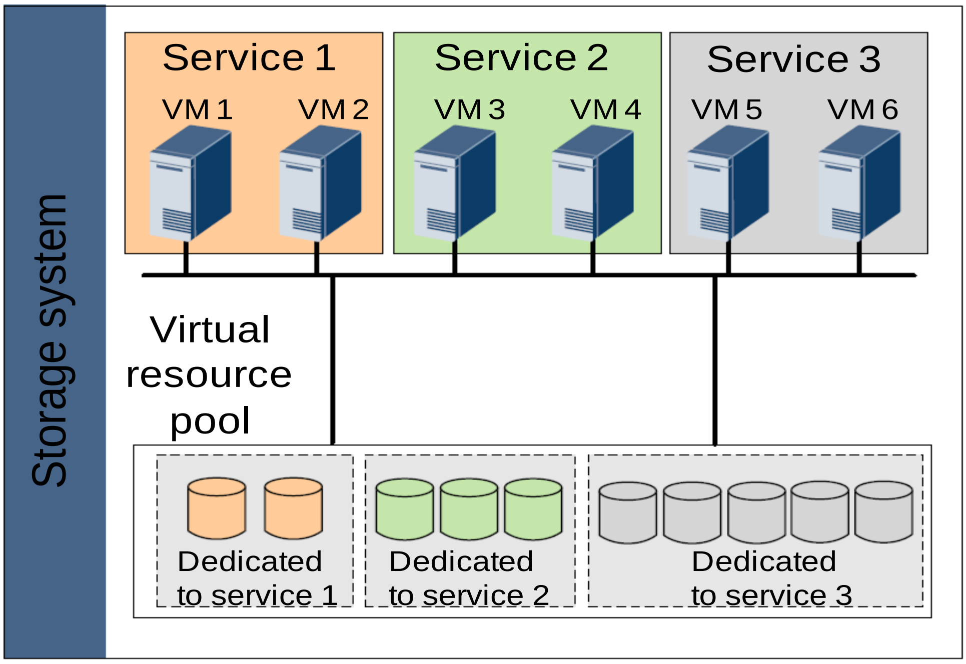
5.4.3 Multi-Service Isolation

Multiple enterprise IT application services run on the same hardware platform for a lower cost and easy maintenance. However, the concurrent service operations require an effective service isolation mechanism to ensure service stability, reliability, and security.

SmartContainer uses VMs to isolate applications of multiple services.

1. Mainstream operating systems and their applications can be deployed on different VMs for isolation.
2. Operating systems and applications on VMs are independent from each other and VMs use their own storage resources. Services and applications are securely isolated.

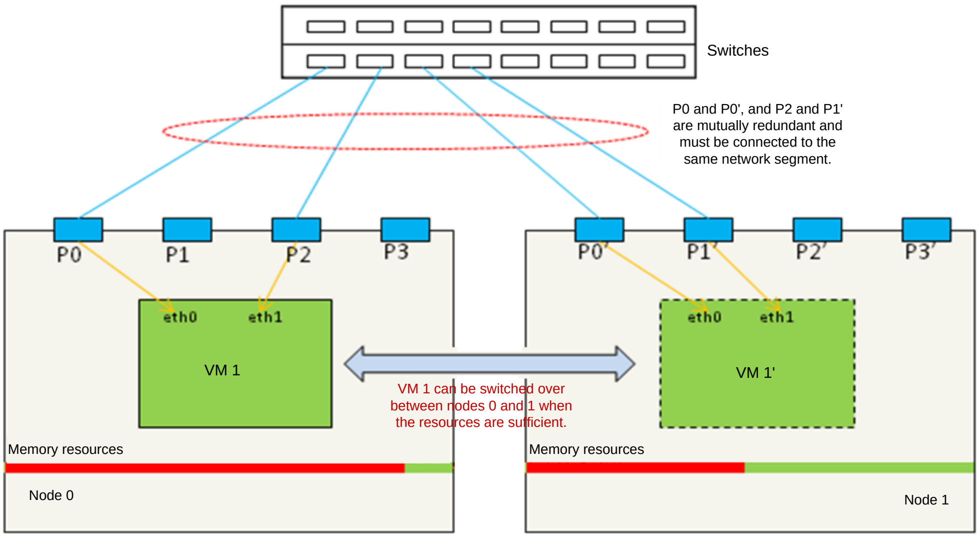
The following figure shows a multi-service isolation scenario.



5.4.4 High Availability of VMs

SmartContainer supports controller failover. As resources on the standby controller are limited and importance of services running on each VM may be different, a high availability (HA) switch is configured for each VM, allowing you to determine whether to migrate the VMs to the standby controller if the active controller fails.

A series of resource check operations must be performed before the HA switch is enabled. The HA switch can be enabled for a VM only when the VM has sufficient image resources.



5.4.5 Flexible Enabling Policies

SmartContainer can be flexibly enabled or disabled based on service requirements. After this feature is disabled, CPU and memory resources allocated to VMs will be reclaimed and returned to the storage system to maximize storage system performance. You need to restart the system after enabling or disabling this feature.

6 Conclusion

The OceanStor 2800 V5 video surveillance storage system produced by Huawei is dedicated to the video surveillance field, with features such as low price, excellent performance, large capacity, high reliability, flexible scalability, and easy maintenance. Truly it can help customers lower the TCO of video surveillance applications, simplify management, and improve performance.

7 Acronyms and Abbreviations

Acronym or Abbreviation	Full Spelling	Explanation
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	
VM	Virtual Machine	
VMFS	Virtual Machine File System	A VMFS is used to store ISO images of system disks and operating systems on VMs, VM templates, and VMTools.
VMContainer	Virtual machine container	A resource pool used to run VMs

Acronym or Abbreviation	Full Spelling	Explanation
QEMU	NA	An open-source VM simulation environment component
KVM	Kernel-based Virtual Machine	
VMTools	UVP VMTools	UVP VMTools is a management tool used for optimizing VM I/O performance and easing VM management. UVP VMTools can enhance I/O processing efficiency of host VMs and can allow upper-layer services on VMs to communicate with hosts. To better manage VMs and enhance the working efficiency of VMs, it is required that UVP VMTools be installed on all VMs.
VHBA	Virtual Host Bus Adapter	
vHost	vHost	A new type of initiator driver similar to the FC and ISCSI at the storage layer. vHost allows LUNs at the storage side to be presented on Euler OS as block devices.