Huawei FusionCube Database Infrastructure Technical Proposal

Issue 01

Date 2018-02-01

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://e.huawei.com

Email: support@huawei.com

Table of Contents

1 Project Overview

1.1 Project Background

//(Delete this paragraph before delivering the document to customers.) The information described in this section is for reference only. Describe the actual background of your project.

With the information boost following technical innovations, the XX company needs to analyze massive information from multiple perspectives to obtain the information useful to decision-making. The transaction-oriented operational databases are becoming inadequate to meet the new requirements. Currently, data warehouses have become a core business intelligence technology that can convert data into knowledge and ultimately offer profits for enterprises. However, the current data warehouse used in the XX company confronts the following challenges:

- **Large data volume**

  Continuous development of services causes the explosive growth of data, thereby imposing high requirements for the capacity and expansibility of data warehouses. The capacity of data warehouses is required to be scalable to hundreds of TB or dozens of PB.

- **Various data categories**

  Besides the data generated in traditional transaction systems, now more and more text data, images, and voice data also need to be poured into data warehouses for analysis and processing. Data warehouses need to be able to process both structured and unstructured data.

- **Hybrid workloads**

  Traditionally, transaction data is processed in online transaction processing (OLTP) databases. Now more and more new services require real-time processing. Therefore, data warehouses are required to be able to process data produced in both OLTP and online analytical processing (OLAP) databases.

- **High data accuracy requirements**

  Data in a data warehouse may be from different source databases and of various data types. Data searches must support flexible combination of search criteria. Moreover, enterprises are increasingly dependent on database search results to make business decisions. Therefore, data warehouses are required to present better search consistency and accuracy performance.

- **High network bandwidth requirements**

  With the increase of data categories and data sources, more and more data needs to be processed, making the network bandwidth required increase from hundreds of Mbit/s to dozens of Gbit/s. Therefore, the traditional network architecture can no longer meet service bandwidth requirements.

- **High data query speed requirements**

  Due to the large data volume, various data types, and high search complexity, search services have a common challenge arisen for enterprises. This results in slow search responses, which may take several to dozens of minutes or even fail to return any results at all.

To conquer the challenge, the XX company need to build a new data warehouse system or upgrade the existing data warehouse system based on the customer's live network conditions, service requirements, and future service plan. After the data warehouse system is built or upgraded, the customer can migrate its data stored on separated database systems to the data warehouse in the extract, transform, and load (ETL) mode.

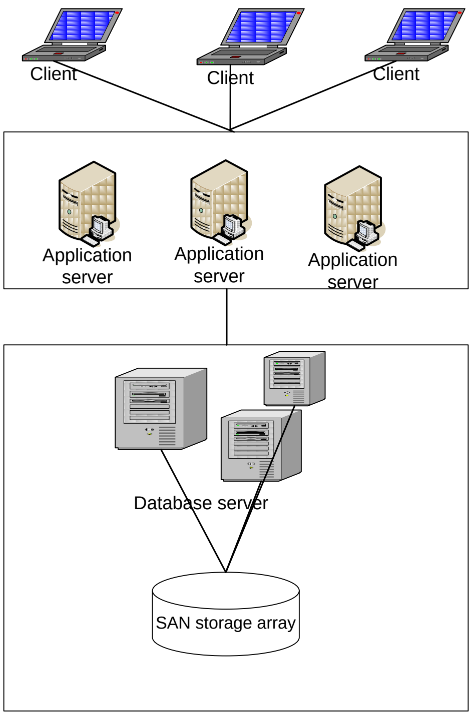
1.2 (Optional) Traditional Architecture and Challenges

//(Delete this paragraph before delivering the document to customers.) The information described in this section is only an example for reference. Modify the content in this section based on actual conditions.

1.2.1 Traditional Solution Architecture

If the traditional architecture is used to deploy the information system for the XX company, the logical architecture of the system is as follows:

- Computing: One server is used, or multiple servers are clustered with the database management software installed.
- Storage: The traditional storage area network (SAN) arrays are used to store user data.
- Network: All servers are connected to SAN arrays using Ethernet or Fibre Channel (FC) switches.
- High availability: The backup server and disk arrays are deployed, and the data replication software is used to back up data to a backup server or a local directory.



1.2.2 Challenges for the Traditional Solution

The traditional solution has the following shortcomings:

- Computing, storage, and network devices are separately deployed, making hardware installation and deployment complex.
- Different devices must be maintained on different systems, and no unified monitoring and alarm reporting mechanisms are available, incurring high management and maintenance costs.
- Upgrade and capacity expansion of servers, mid-range computers, and storage devices are complicated and costly.
- Computing, network, and storage are separate from one another, resulting in bandwidth, I/O latency, disk I/O, and computing capability bottlenecks during resource access. The system performance bottlenecks cannot be eliminated by only upgrading servers, mid-range computers, storage devices, or network devices.

1.3 (Optional) Background of Midrange Computer Elimination

//(Delete this paragraph before delivering the document to customers.) The information described in this section is only an example for reference. Modify the content in this section based on actual conditions.

With the rapid development of x86 platforms, many industries, such as the telecom industry, public security sector, healthcare institution, and financial industry, plan to replace mid-range computers with x86 platforms and migrate their services to x86 platforms to consolidate their IT systems.

Armed with Intel® Xeon® E5 v4/E7 v4/Skylake series processors, the x86 platform support the Intel Turbo Boost, hyperthread, virtualization, and intelligent energy saving technologies. x86-based servers using the Xeon E7 series processors can provide the same performance and even outperform RISC-based mid-range computers with only 20% to 50% costs. For the most important RAS features in the mission-critical service field, according to Intel, the Xeon E7 series processors can provide similar reliability, availability, and serviceability with Intel® Itanium® processors. This allows server vendors to develop x86 systems with an availability up to 99.99999%, 99.999999%, or even 99.9999999%.

To enable a midrange computer to provide high stability and scalability, the hardware and software of the mid-range computer are better from the same vendor. If third-party hardware and operating systems (OSs) are used, vendors of mid-range computers must test the software compatibility to ensure the system stability, which is time-consuming. However, for x86 systems, third-party vendors are required to test the compatibility of hardware and OSs by themselves, which helps lower the costs of x86 systems. In the systems or solutions where mid-range computers are used, the processors, platforms, OSs, middleware, and application software are all from the mid-range computer vendors. The advantage is that the hardware and software in a mid-range computer can be integrated and optimized to provide one-stop purchase, deployment, and maintenance, freeing users from heavy hardware and software selection and system maintenance workloads. The disadvantage is that this mode will cause vendor lock-in, making IT system device procurement, maintenance, and application monopolized by one vendor. In additions, this mode creates barriers for users when migrating services from the mid-range computer-based system to other systems.

Therefore, the computing platform requirements must be considered at the beginning of service system construction. Service systems must be deployed on computing platforms running common OSs and support distributed cluster deployment mode to maximize return on investment (ROI) and minimize total cost of ownership (TCO). Because x86 platforms can meet all the preceding requirements, they are undoubtedly the best choice for users to deploy their service systems. The hardware cost of x86 platforms is only a fraction of that of mid-range computers.

1.4 Project Requirements

//(Delete this paragraph before delivering the document to customers.) Modify the content in the following table based on actual project requirements.

Table 1-1 Project requirements

No.	Item	Requirements	Remarks
1	Computing capacity	A total of 96 CPU cores are required.	Both 2-core and 4-core CPUs can be used.
2	Storage capacity	The system must use SAS disks and provide 30 TB effective storage capacity.	
3	Performance	The system must support more than 6 million transactions per minute C (tpmC) and 1000 concurrent users.	
4	Backup	A backup system must be built at the local site. The backup policy is set to allow the system to back up all data every one week and incremental data every seven days and to retain full-backup data generated in one month. The total capacity required for backup is 60 TB.	Optional
5	Disaster recovery (DR)	Active-active DR and third-place DR are required for database applications.	Optional

2 Solution Overview

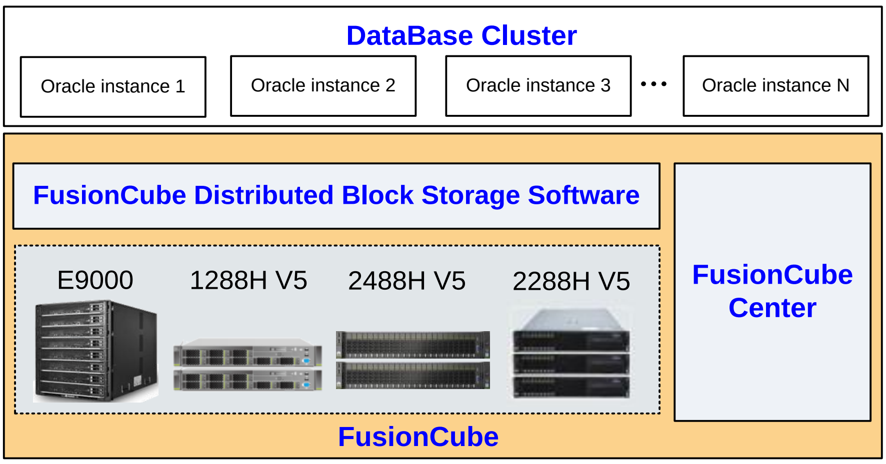
This project uses the FusionCube database infrastructure platform.

2.1 Architecture of the Database Infrastructure

//(Delete this paragraph before delivering the document to customers.) The information described in this section is for reference only. Describe the actual architecture for your project based on customer requirements.

FusionCube is Huawei's proprietary infrastructure that integrates computing, storage, and network resources on a converged platform.

Figure 2-1 Huawei FusionCube architecture



It consists of FusionCube Center, FusionStorage Block software, and a hardware platform.

FusionCube Center: It features comprehensive management such as resource management and operation maintainance for FusionCube, providing a unified web-based user interface (WebUI) for centralized resource management, service display, and O&M, including user management, hardware monitoring, alarm reporting, and storage resource configuration.

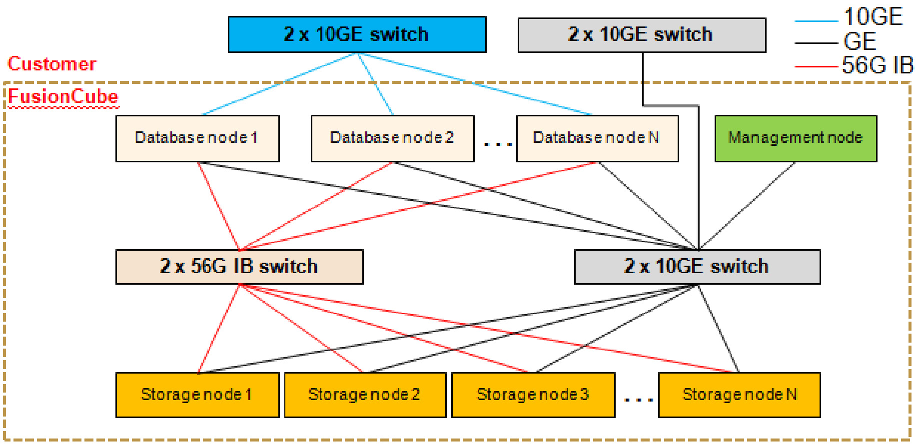
FusionStorage distributed block storage software: It adopts distributed cloud storage technology to virtualize the local hard disks of servers into a shared storage pool. It employs a scale-out architecture that allows for horizontal scalability and provides block storage services with high performance and reliability.

The distributed block storage system of FusionCube can effectively solve the problems caused by SAN storage. It adopts the new-generation distributed storage architecture and parallel distributed grid storage technologies. The scale-out architecture and distributed multiple-node grid implement storage load balancing. Fine-grained data distribution algorithms are used to ensure even data distribution, improving system reliability.

Hardware platform: FusionCube supports converged blade servers and general-purpose rack servers, and supports the latest Skylake servers.

2.2 Logical Network Architecture

//(Delete this paragraph before delivering the document to customers.) The information described in this section is for reference only. Describe the actual architecture for your project based on customer requirements.



The entire network is divided into two layers:

1. Access layer

   Connect servers and storage devices uplink to switches at the access layer.

   Deploy six NICs on the server, two for the service plane, management plane, and storage plane, respectively, to ensure link redundancy.

   Add the management, service, and storage planes to VLANs on the access switches to isolate them. You are advised to stack switches to simplify networking and improve network reliability.

   Service plane network: 10GE network, which transmits service data between compute nodes.

   Storage plane network: 56 Gbit/s InfiniBand network, which transmits service data between compute nodes and storage nodes.

   Management plane network: GE network, which transmits internal management data between compute nodes and storage nodes.

2. Aggregation layer

   Connect access switches uplink to the switches at the convergence layer. You are advised to configure aggregation switches to work in cluster mode. Connect access switches to aggregation switches through Eth-Trunk ports. After the aggregation switches are stacked, the VRRP function is not required. If aggregation switches are required to provide the gateway function, set the user gateway to the IP address of the VLANIF interface.

2.3 Detailed Configuration

Component	**Specifications**	**Quantity**
Cabinet	42U, 19-inch standard cabinet	1
Management node	1288H V5: 2 x Xeon Silver 4110, 4 x 16 GB memory, 2 x 600 GB SAS, 2 x GE, 2 x 10GE, 1 x LSI3008	1
Database node	1288H V5: 2 x Xeon Platinum 8160, 16 x 32 GB memory, 2 x 600 GB SAS, 2 x GE, 4 x 10GE, 1 x Avago3408, 2 x 56 Gbit/s IB	2
Storage node	2288H V5: 2 x Xeon Silver 4114, 5 x 32 GB memory, 2 x 600 GB SAS, 4 x 3200 GB NVMe SSD, 12 x 8 TB SATA, 2 x GE, 2 x 10GE, 1 x Avago3408, 2 x 56 Gbit/s IB	3
Management switch	Huawei CE5855: 48 x GE electrical ports	2
InfiniBand switch	MELLANOX: 36 x 56 Gbit/s InfiniBand ports	2
Storage software	FusionCube distributed storage software standard edition, 3-year subscription and support service (per TB)	288

3 Solution Highlights

3.1 High Reliability

3.1.1 Server Reliability

Computing server reliability includes HA and reliability of memory, hard disk, and power.

- The computing server hardware provides the following basic reliability capabilities:
  - Basic input and output system (BIOS) memory self-check and error checking and correcting (ECC) technologies
  - Hard disk hot swapping, RAID, and in-service hard disk fault detection and warning
  - Power supply 1+1 redundancy and hot swapping
  - Real-time temperature monitoring and alarming of heat-sensitive components, such as CPUs, DIMMs, fans, PSUs, and hard disks. The server supports replacement of hot-sawppable components, and power-off replacement of the components which are not hot swappable. Intelligent fan speed controlling and monitoring improve system reliability.
- The computing server provides the HA feature:
  - that isolates a faulty computing server from its RAC cluster using RAC software while keep original services running in the cluster.

3.1.2 Storage Reliability

- Link redundancy

  In the computing-storage-separated architecture, at least two redundant paths need to be configured between each compute node and storage node so that the system can implement the multipathing access function of the storage. The software automatically performs switchover between multiple paths after a fault occurs. This function prevents storage access failures due to faults on a single node.
- Storage data backup

  FusionStorage Block implement data copies. When a storage node is faulty or some hard disks on a storage node are faulty, the system automatically rebuilds data without affecting services.

3.1.3 Network Reliability

- Redundant network paths

Core switches use CSS technology to ensure redundancy of external connections to firewalls or the network address translation (NAT) and internal connections to aggregation switches.

Aggregation switches use switch cluster technologies to provide redundant connections to firewalls (uplink) and to aggregation switches (downlink).

Access switches use stacking technology to ensure redundancy of uplink connections to aggregation switches.

- Plane-based network communication

  The system uses three planes for communication: management plane, storage plane, and service plane. To ensure data reliability of all network planes, these planes are separated by VLANs. If one plane malfunctions, the other two planes still keep working properly. Service plane: An internal InfiniBand or 10GE network for communication between compute nodes.

  Storage plane: It is a communication plane between compute nodes and storage nodes, which is provided by an internal InfiniBand or 10GE network.

  Management plane: It processes services including node management, service deployment, and system loading for all nodes in the database.

- NIC load balancing

  Two NICs are provided for each communication plane (service, storage, and management) and are bound as one logical NIC to function in bonding mode. This mode balances server traffic load among NICs and ensures service continuity by switching over services to another NIC if one NIC becomes faulty.

3.1.4 Management Reliability

- Separation of computing and storage clusters

  By default, computing and storage clusters are separated to improve system reliability. The computing clusters distribute query tasks of users and aggregates results. The storage cluster performs data I/O operations on disks. A compute node failure does not affect the proper running of the storage cluster, and a storage node failure does not affect the proper running of the computing cluster.

- Fault detection

  The fault detection function supports server, software, and resource monitoring. By enabling the detection function on each monitored node, the system can collect key server performance indicators, such as the CPU usage, infrastructure network traffic, and memory data. Faults, such as process exceptions, management and storage link exceptions, node exceptions, and system resource overload, can be detected. This improves the fault detection capability of the system.

  The fault detection function supports fault information collection and availability measurement of storage cluster nodes. This information can be displayed on the web browser. Users can view information about cluster management and system load allocation, and determine whether problems, such as load balancing problems, out-of-control processes, and hardware performance deterioration, exist. This function helps adjust system resources and improves system performance. Historical records of hardware resource usage can be viewed by day, week, or year.

- Black box

  Management and compute nodes use black box technologies to improve fault addressing capabilities. If a system exception occurs, the system automatically stores its kernel logs, system snapshots, kernel diagnosis, and last words on a secure storage device, such as a compute node, or uploads the information to a network server, such as a log server, to allow users to analyze and identify fault causes.

3.2 High Security

3.2.1 Security Structure

To ensure security of data centers, the Huawei FusionCube Database Infrastructure adopts complete security architecture to enhance network isolation and virtualization isolation. This security architecture follows the ideas of layered defense and in-depth defense.

Layered defense: Layered defense employs multiple methods to implement security policies in various network areas to ensure that no single point of failure (SPOF) occurs on the network.

In-depth defense: In-depth defense uses multiple security policies to manage risk. If one security policy fails, another policy takes effect to prevent devastating damage.

Huawei security framework is divided into several layers, physical layer, host layer, network layer, service and data layer, and O&M layer, based on the network architecture to ensure security in the following aspects:

- **Physical device security**

  Only authorized personnel can enter the equipment room where the data warehouse FusionCube infrastructure is located. Moreover, monitoring devices are installed in the equipment room for viewing site conditions if necessary.

- **Infrastructure security**

  The FusionCube Database Infrastructure has various BI applications installed and allows OLTP data insert operations. Therefore, it is prone to security attacks from sources including viruses, vulnerabilities, Trojan programs, and denial of service (DoS) attacks. The system ensures infrastructure security by system hardening, antivirus software, and security patches.

- **Network security**

  The data warehouse network may be attacked by a variety of security risks, such as DoS attacks, eavesdropping and tampering of user data. Huawei FusionCube ensures network security using various methods including network isolation, attack prevention, and transmission security assurance.

  Network partition and isolation isolates the computing domain, storage domain, management domain, and access domain. The management plane is an independent physical network. All these measures ensure network security and avoid spreading of network storms.

3.2.2 Network Security

Huawei FusionCube Database Infrastructure ensures database network security using the following protection methods:

- **Security domain partition and network isolation**

  Network partition and isolation separates the computing domain, storage domain, management domain, and access domain. The management plane is an independent physical network. All these measures ensure network security and avoid spreading of network storms.

- **Attack prevention**

  The powerful firewalls constitute the intrusion defense against scanning attacks, malformed packet attacks, resource depletion attacks, and special packet control.

  The firewall uses the NAT function to hide internal networks and isolate security domains for different services, uses access control list (ACL) policies and connection status checking function to check communication validity, and uses the Intrusion Prevention System (IPS) to prevent intrusion at the application layer.

  The built-in IPS module can configure a defense policy based on the destination IP address of the packets, so that traffic with preset IP address segment must pass through the in-depth defense of the IPS.

  Huawei FusionCube uses the anti-distributed DoS (DDoS) function to allow customization of protection policies for large-sized enterprises. Protection method, protected IP address, HTTP port number, and checking thresholds can all be customized based on customer requirements.

  The Secure Sockets Layer (SSL) VPN gateway can be virtualized into multiple virtual gateways to isolate enterprises and implement IP address overlapping.

  In addition, the FusionCube provides a network traffic analysis function to measure traffic of each type of packets in real time. It also performs real-time invasion detection and works with firewalls to report alarms. Moreover, it is able to detect DDoS attacks to prevent attacks from malicious traffic.

- **Transmission security**

  Data transmission may be interrupted, and data may be replicated, modified, forged, intercepted, or monitored during transmission. Therefore, it is necessary to ensure the integrity, confidentiality, and validity of data during data transmission. Huawei FusionCube ensures transmission security using the following functions:

  SSL encryption for data transmitted between trusted zones and non-trusted zones

  Hypertext Transfer Protocol Secure (HTTPS) access for user management services and SSL VPN access for services with high security requirements.

3.2.3 Enterprise Data Security

Data security is critical to data center security. Huawei FusionCube supports the following functions to ensure user data security:

- Residual data deletion

  After a user detaches a disk from a VM to release the disk resource, the system formats the disk before reassigning it to another user. This ensures user data security.

  After user files or objects are deleted from a storage area, the system thoroughly deletes data in the area and marks the area as write-only. This mechanism protects the storage area against unauthorized restoration.

- Data backup

  The system stores one or more backups for each piece of data in data store. This ensures data security and service continuity even if the data storage medium, such as the hard disk, is faulty.
  The system performs bit- or byte-based verification for the stored data and evenly distributes the verification information on each disk in a disk array. Disk arrays save user data blocks and data verification information on different disks. Therefore, if a data disk is faulty, the system can rebuild data on the disk using other data blocks in the same data strip and the verification information.

3.2.4 Management and Maintenance Security

The system supports centralized log collection and storage. It also provides a log auditing system to meet customers' auditing requirements, such as the Sarbanes-Oxley Act (SOX) requirement.

- Log management

  Huawei FusionCube supports the following types of logs:

  1. Operation logs

     Operation logs record the information about operations performed on management nodes. The information, including the operator, operation type, client IP address, key parameters, operation time, and operation result, is saved in the management node databases. By exporting and viewing operation logs, auditing personnel can periodically check for inappropriate or malicious operations performed on management nodes. Operation logs can also be used as operation evidence to address disputes.

  2. Run logs

     Run logs record the running state of each node. The priority of run logs includes debug, info, warning and error in an ascending order. The system can be controlled to generate logs only of a certain level.

     The Rsyslog component collects and filters run logs of each node as high-priority (warning or error) logs or complete logs, including log packages with a set priority. The system periodically transfers higher-level logs to the log server using FTP. Complete logs are stored in a local directory, and the logs generated with a specific duration on a specific node can be uploaded to specified directories on the log server.

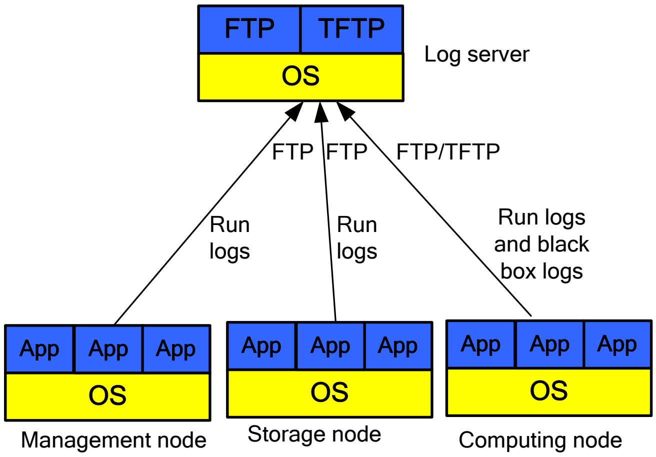
     Run log information includes log level, thread name, and log contents. By viewing run logs, O&M engineers can learn and analyze the running state of the system to detect and handle exceptions in time.

  3. Black box logs

     Black box logs record location information when a critical fault occurs. These logs are used for locating and rectifying faults to recover services rapidly. The system transfers black box logs generated for compute nodes to the log server by using TFTP, and saves black box logs generated for management nodes and storage nodes in local directories.

     Figure 3-1 shows the mechanism of centralized log collection and storage by the system.

Figure 3-1 Centralized log management



An Rsyslog Client is deployed on each node to collect run logs and black box logs generated on the node in real time. High-level logs are exported to specified directories on the log server at a scheduled time. Complete logs are usually stored on local disks. You can run scripts to upload complete logs generated in a specified period and on a specified node to the log server for storage.

- Security alarm management

  After detecting an event or behavior that is against the configured security policies, the system reports alarms to the alarm management system so that the administrator can immediately handle the event to remove security risks. Security alarm information includes alarm source, generation time, alarm cause, service provider, service user, alarm severity, and event type.

3.3 High Scalability

3.3.1 Computing Scalability

The database cluster software provides the expansion capability of compute nodes. Take Oracle database as an example. The Oracle RAC cluster software supports smooth expansion of compute nodes based on a single node and does not affect service running.

3.3.2 Storage Scalability

The FusionCube distributed storage software provides the expansion capability of storage nodes and supports one-click online smooth capacity expansion.

4 (Optional) Backup Solution Design

4.1 Backup Requirements Analysis

//(Delete this paragraph before delivering the document to customers.) The information described in this section is for reference only. Modify this section based on customer requirements.

1. A customer requires a backup system for database and OS data of internal servers in data centers and operation centers. The backup system manages data backups in a centralized manner to improve data security and manageability of each host.
2. Backup software is easy to use and provides wizards for users to configure backup and restoration policies for applications, files, and Structured Query Language (SQL) databases. In addition, the backup software needs to support calendar-based task management, facilitating query of daily backup tasks.
3. The data of XX service hosts in the XX service system must be protected. Table 4-1 lists data volume of each service system.

Table 4-1 Backup requirements analysis

System Name	Service Host Name	Backup Object	Initial Backup Data Volume (GB)	Estimated Incremental Data Volume (GB)
OCS	*LIGDB*	*DB*	*100*	*1.4*
	Report	*DB*	*100*	*0.7*

The backup system is required to:

- Finish all backup tasks in XX hours.
- Provide a trial function for users to make verification and modification for the configured backup policies.
- Back up and restore database, file, and OS data in a quick and effective manner.
- Allow the backup management software to control resource use in backup and recovery operations as required. For example, the backup system allows assignment of server network ports for backup and recovery traffic transmission to reduce network bandwidth usage, and also allows customization of firewall ports.
- Provide reliable backup software.

- Consider system expansion requirements during backup system design to ensure smooth system upgrades.

4.2 Guidelines for Backup System Design

Data backup ensures that data can be recovered in case of a system failure or disaster, thereby relieving users and operators' worries about data loss. Backup solutions vary depending on different application scenarios. Typically, a well-defined backup system must comply with the following performance requirements:

- Stability

  A backup system is used to provide data protection for the system. Therefore, the backup system must be stable and reliable. This requires that the backup system must be able to be completely compatible with the OS and to recover data in a quick and effective manner when a fault occurs.
- Automation

  Many systems have requirements on the backup start time and backup window. Backup operations are usually recommended to be performed at midnight when the service load is light, which, however, increases the workload of system administrators. Therefore, the backup solution must be able to provide automatic backup functions and automatically manage backup media equipment. During the backup, the backup system must be able to record logs and generate alarms to report exceptions and faults.
- High performance

  An increasing amount of data is generated and updated frequently to meet the increasing service demand. If data backup cannot be complete during the off-hour, the data backup performed during working hours affects system performance. This dilemma requires taking the data backup speed into consideration when planning backups to ensure that data backup can be accomplished within the specified time period.
- Service system validity

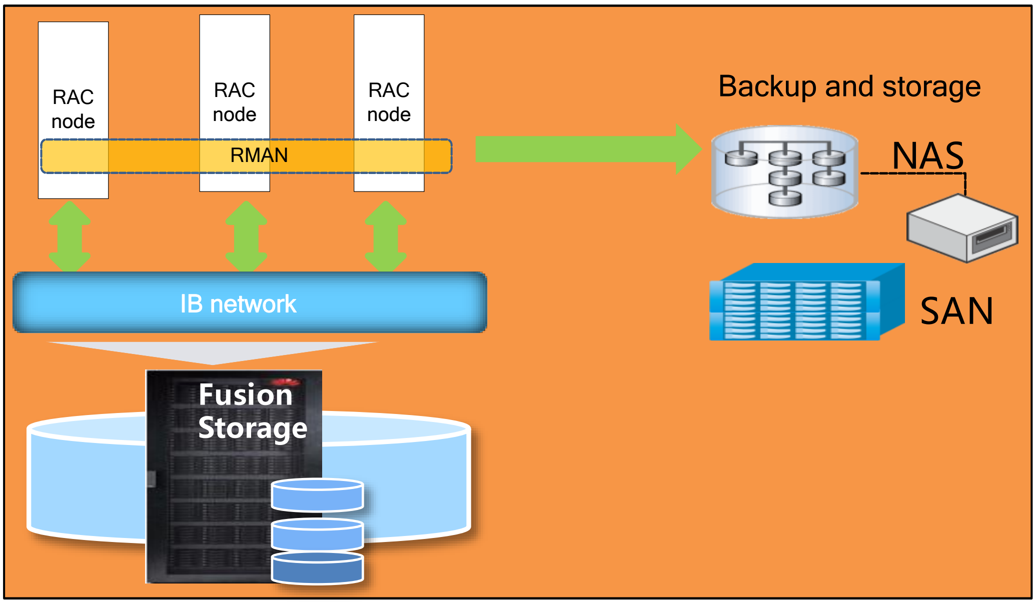
  Data backup may affect service system performance. It is important to use effective technical measures to prevent the server system, database system, and network system from being affected by data backup. In addition, data that is restored by the backup system must be valid and complete.
- Simple operation

  Data backup is used in different areas. Operators who perform data backup belong to different levels. An intuitive and simple graphical user interface (GUI) is required to shorten operators' learning curve, reduce their working pressure, and facilitate data backup configuration.
- Real-time backup performance

  Some key tasks require 24 hour uninterrupted operation. Some files of these tasks may still be open during data backup. In this case, measures are implemented to check file size in real time and perform event trace to ensure that all system files have been correctly backed up.

4.3 RMAN Backup Solution

Figure 4-1 RMAN backup solution



RMAN is a backup and recovery manager developed by Oracle. It supports backup on various backup devices, such as storage area network (SAN), network attached storage (NAS), and magnetic tapes. RMAN is able to back up diverse objects including control files, tables, and schemas, and also supports both full backup and incremental backup. Backup policies can provide full or incremental backups based on service requirements.

4.4 Backup Policy Design

The priorities of data to be backed up for the XXX service system are as follows:

- Priority 1: database data
- Priority 2: important configuration files and log files

Backup policies for data with different priorities are as follows:

Priority 1: user data and service data saved in databases

Backup policy: Implement a full backup every day and an incremental backup every 4 hours for the production system. Retain the backup data generated in recent 28 days, that is, retain the data generated in 29 backup cycles (28/1 + 1 = 29). Implement a full backup every week and an incremental backup every day for the test system. Retain the backup data generated in recent 28 days, that is, retain the data generated in five backup cycles (28/7 + 1 = 5).

Table 4-1 Backup policy plan

System	Backup Object	Backup cycle	Retention Period	Backup Type	Backup Mode
Production system	Database data	Full backup: 1 day Incremental backup: 4 hours	28 days	Online backup	Full and incremental backup
Test system	Database data	Full backup: 7 days Incremental backup: 1 day	28 days	Online backup	Full and incremental backup

4.5 Backup Capacity Design

//(Delete this paragraph before delivering the document to customers.) Values listed in the table in this section are for reference only. Enter values based on actual project information.

Backup media capacity = (Initial backup data volume + Full backup cycle/Incremental backup cycle x Estimated incremental data volume) x (Retention cycle/Full backup cycle + 1) = (A + C/D x B) x (E/C + 1)

Initial backup data volume: A GB

Estimated incremental data volume: B GB

Full backup cycle: C days

Incremental backup cycle: D days

Retention period: E days

Table 4-1 Backup volume plans

System Name	Service Host Name	Backup Object	Initial Backup Data Volume (GB)	Estimated Incremental Data Volume (GB)	Full Backup Cycle (Day)	Incremental Backup Cycle (Day)	Retention Period (Day)	Backup Media Volume (GB)
OCS	*LIGDB*	*DB*	*200*	*1.4*	*1*	*1/6*	*28*	*6044*
	Report	*DB*	*100*	*0.7*	*1*	*1/6*	*28*	*3022*

4.6 Backup Window Design

Information about the backup network is as follows:

The initial data volume is estimated to be 1000 GB, and incremental data volume per day is 100 MB.

The available bandwidth of the backup network is 1000 Mbit/s.

Full backup window for the first time:

Assume that the available network bandwidth is 1000 Mbit/s, the bandwidth usage is 75%, and the volume of the data to be fully backed up is 200 GB. The first full backup window is calculated as follows:

(1000 x 1024 x 8)/(1000 x 75%)/3600 ≈ XX (hours)

The incremental backup window is calculated as follows:

(100 x 8)/(1000 x 75%)/3600 ≈ X (hours)

4.7 RMAN Backup Advantages

RMAN provides effective protection for user data. The advantages are as follows:

- Ease of use: The NAS or SAN devices can be attached to a directory of the RAC node. The database data can be directly backed up to a backup storage device by running RMAN scripts or backed up to the local storage space of the RAC node and then stored to a large-capacity storage device such as taps.
- Flexible backup policy configuration: RMAN supports periodical full and incremental backup. Users can configure a backup cycle, backup window, data backup retention period, and obsolete data backup deletion policies. Users can also allocate backup channels for data backup.
- Restoration: Users can select restoration objects based on service requirements and select proper restoration channels based on the volume of data to be restored.

5 (Optional) DR Solution Design

5.1 DR Requirement Analysis

//(Delete this paragraph before delivering the document to customers.) The information described in this section is for reference only. Describe the actual DR requirements for your project based on customer requirements.

Configure a DR system for important service systems to ensure operation continuity. Table 5-1 lists the requirements for configuring a DR system.

Table 5-1 Disaster recovery (DR) requirement

No.	System Name	Server Name	System Type	DR Directory	Current Data Volume (GB)	Snapshot Frequency (Minute)	Copies of Retained Snapshots	RTO (Minute)	RPO (Minute)	Remarks
1	*Production system B*	*XXX database*	*Database*	*/oracle*	*20*	*30*	*30*	*30*	*30*	

5.2 DR Solution Design

5.2.1 DataGuard Solution

Oracle Data Guard is one of the most efficient enterprise data protection solutions in the industry. It is capable of protecting data against loss all around the clock in case of any disaster or communication interruption.

Oracle Data Guard provides the management, monitoring, and automation software to create, maintain, and monitor one or more standby databases to protect enterprise data from failures, disasters, human errors, and data corruptions.

Data Guard stores a database backup that contains the same transaction data as the production database. The backup databases may be located in a DR site that is in the same building, campus, or city as the production center, or may be thousands of miles away from the production center. If the production database becomes unavailable due to planned downtime or unexpected interruptions, Data Guard is able to switch any backup database to the production database. This mechanism minimizes the service downtime and prevents data from being lost.

Oracle GoldenGate is a software product for structured log data replication. It obtains changes made to data in the source database by parsing online or archived logs, and synchronizes the changes to the target database in active-active mode.

The real-time data replication technologies used by GoldenGate ensures data security while preventing application of a variety of software and hardware products. It simplifies O&M complexity and investment and facilitates DR system switchover and recovery.

5.2.2 DR Switchover

Generally, a DR system performs switchover in two conditions. One is normal switchover that occurs in a DR switchover practice or switchback to the production system. The other condition is DR switchover that is triggered when the production center is faulty.

The DR center duplicates data from the production center in an asynchronous manner. If any data error occurs due to human misoperation, the data in both the DR center and the production center will become incorrect. In this case, you can use the local historical snapshot data to restore the system to the status before the exception occurs. Besides data, DR systems play greater importance in ensuring service continuity.

Table 5-2 describes the handling suggestions for some common failures.

Table 5-2 DR switchover policies

Fault Cause	DR Switchover	Handling Suggestion
A single device is faulty.	*No*	*Use the redundancy system of the production center to ensure service availability and fix the faulty device as soon as possible.*
The storage system of the production center experiences a major fault and cannot be rectified in a short period.	*Yes*	*Use the DR center as the active system to provide services.*

Fault Cause	DR Switchover	Handling Suggestion
Applications are faulty and cannot be restored in a short period.	*Yes*	*Use the DR center as the primary system to provide services.*
The site where the production center is located encounters a disaster, causing a breakdown of the production center.	*Yes*	*Use the DR center as the primary system to provide services.*

5.1 Advantages and Features of the Disaster Recovery Solution

The DR solution brings the following advantages:

- Reduced investment

  Oracle Data Guard is delivered with Oracle database products for free. Therefore, users do not need to pay extra fees on DR software, reducing their investment on database software.

- Mature and stable technologies

  Oracle has already launched Data Guard functions in the Oracle 7 edition, then named Standby database. Using Oracle';s mature archiving, backup, and restoration technologies, Data Guard now is recognized as a stable DR software that has won a wide range of successful application cases.

- Little impact on system running

  Data Guard does not parse logs on the primary database server. Instead, it only requires the primary database server to transfer the archived log files to the DR node. Therefore, the DR implementation imposes little adverse impact on the performance of the production center.

- Meeting users' basic service requirements

  Data Guard is capable of meeting basic service requirements, such as data DR, RTO, RPO, and transmission bandwidth.

The DR solution has the following features:

- Real-time DR performance: After a new transaction is produced at the data source, the DR system immediately captures the data, converts it if needed, and transfers the data to the destination system within a short period time.
- High availability: GoldenGate ensures high service availability. Services in the system are not interrupted even in case of a power outage or a planned power-off.
- Support for heterogeneous devices: If only the source and destination sites use mainstream databases and platforms, the TDM can be used to duplicate data from each other even heterogeneous devices are used in the two sites. This feature facilitates flexible server configurations for customers.

- High performance and low impact: GoldenGate is capable of handling thousands of transactions per second while hardly imposing adverse impact on system performance.
- Transaction consistency: GoldenGate ensures transaction integrity and consistency although the source and destination sites use different systems.

5.2 DR System Configuration Design and Device List

Assume that the service-busy duration in the production center is 8 hours, the compression rate is 4, and the bandwidth usage is 0.7, the service system data traffic (xxx) is calculated using the following formula: xxx = Number of database instances x Daily service data throughput (GB) x 8/(8 x 3600)/4/0.7. Based on the calculated service data traffic, two xxx Gbit/s links are required between the primary and redundancy databases to ensure service continuity when any single link is faulty.

//(Delete this paragraph before delivering the document to customers.) The information described in this section is for reference only. Describe the actual DR requirements for your project based on customer requirements.

A total of XXX DR devices are deployed based on customers' requirements. XXX DR devices are deployed in the XXX data center, and XXX DR devices are deployed in the XXX data center.

Table 5-3 DR devices

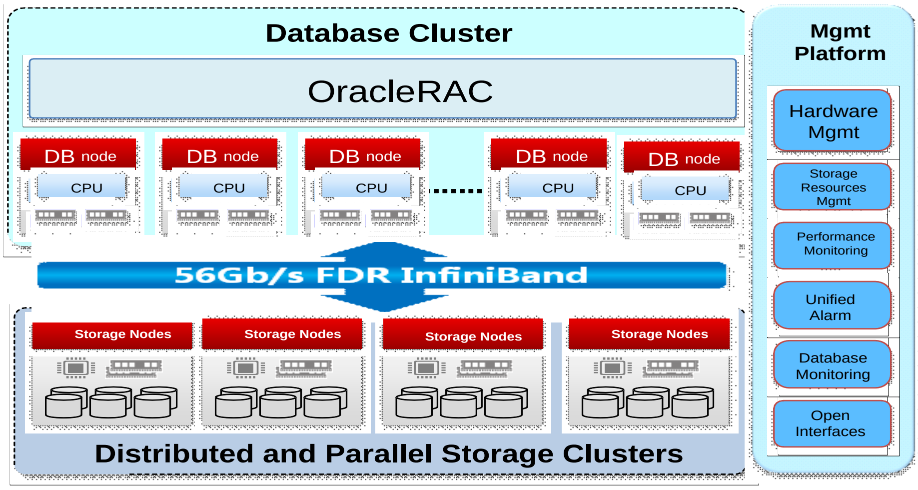
No.	DR Device	Location	Disk Space (GB)	Quantity	Remarks
1					

6 FusionCube Database Infrastructure

6.1 System Architecture

Figure 6-1 shows the overall architecture of the FusionCube database appliance.

Figure 6-1 FusionCube overall architecture



Item	Introduction
Blade/rack server	Hardware component that provides physical computing, and storage resources for FusionCube. It must be used together with the FusionStorage Block software. Computing and storage nodes are configured on demand. SAS disks and PCIe SSD cards are supported. 10GE and InfiniBand switching networks are supported.
FusionStorage Block software	A software component that manages disks on storage nodes and provides storage resources for FusionCube.
FusionCube Center	A software component that manages FusionCube virtualization and hardware resources and implements system monitoring, O&M, and service log management.

6.1.2 Unified Management

The centralized O&M system of FusionCube is implemented by FusionCube Center. FusionCube Center manages hardware devices and resources pools. It also provides a unified portal for the lifecycle management of various resources in the system.

FusionCube Center provides the following functions:

Unified resource management: FusionCube Center aggregates the resources of different systems and establishes a unified resource management model to provide unified resource management, such as monitoring and alarm reporting on various resources.

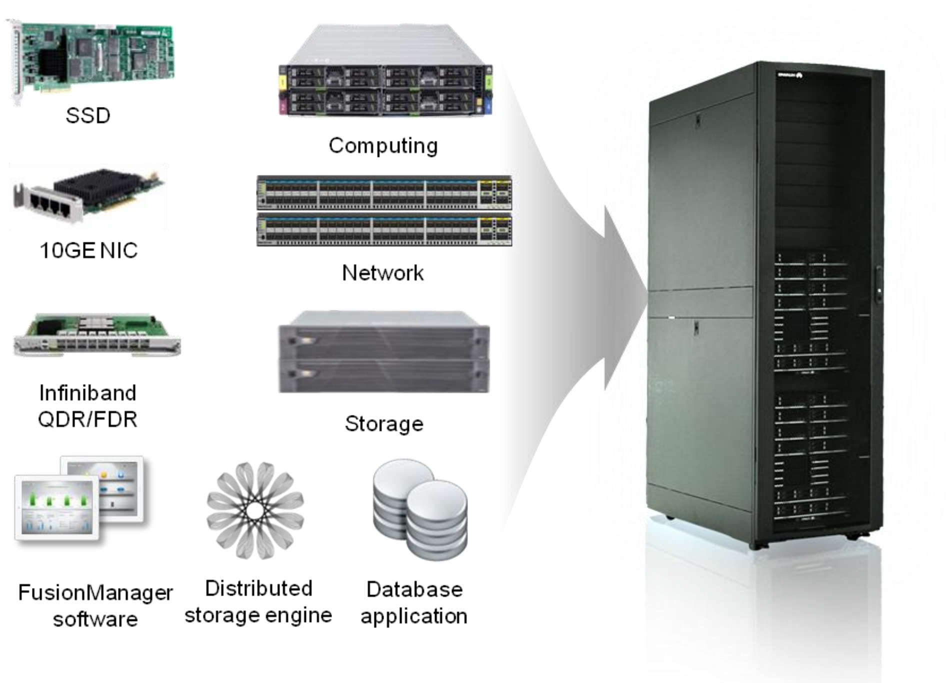
Role-based Access Control (RBAC): RBAC is an effective access-control mode that applies to enterprise security management. RBAC assigns the rights to perform specified operations to specific roles. Each role has a group of rights. Once a role is defined for a user, the user can perform all the operations authorized to the role. With RBAC, you only need to allocate a role to a new user without allocating rights. Because role right changes are fewer than user right changes, RBAC simplifies user right management and reduces system costs.

Automatic hardware discovery and capacity expansion: The FusionCube system can automatically discover all hardware in the appliance and scan the hardware specification information. When creating a resource pool, the user can select the hardware discovered by the system, and then the system will automatically complete software installation and network configuration for the hardware, and add the hardware to the resource pool. During system capacity expansion, the system can also automatically discover the newly added hardware, and automate all configuration jobs when the user adds the hardware to a resource pool.

6.1.3 In-depth Convergence

Figure 6-9 illustrates the converged hardware forms of the FusionCube DB database infrastructure:

Figure 6-2 FusionCube converged platform



Huawei FusionCube database appliance adopts industry-leading modular components and is available in diversified combinations of product models. One-stop services is provided for FusionCube, meeting the requirements of various service ranges. Huawei FusionCube is built on Huawei computing, storage, and network devices, and employs the converged hardware platform architecture with vertically integrated computing, storage, and networking. It features high-bandwidth, low-latency multi-protocol switching capabilities, thereby boosting application performance. As an in-depth integration of software and hardware, Huawei FusionCube defines the very true sense of a hyper-converged appliance.

Single-pane-of-glass: Users perform all operations, such as configuration, alarm management, and performance statistics collection, related to computing, storage, and network resources through a unified user interface, ensuring consistent user experience.

One-stop services: Huawei provides E2E one-stop support for hardware, software, and applications, ensuring that users can obtain optimal services from Huawei.

6.1.4 Elastic Scaling

FusionCube supports elastic capacity expansion based on service requirements without interrupting running services.

If servers in a cluster are insufficient, users can add computing server nodes to the cluster.

If the storage resources in a cluster are insufficient, users can add storage server blades or disks to the cluster.

FusionCube supports linear cluster capacity expansion without service interruption. For each cluster, the computing and storage resources are independent, but the O&M and service provisioning are implemented through one FusionCube Center system in a centralized manner.

6.1.5 Pre-integration

Huawei FusionCube has all its components installed before delivery, thereby simplifying onsite installation and commissioning, and shortening the deployment time from several weeks or even months to only a few days.

FusionCube pre-integration includes:

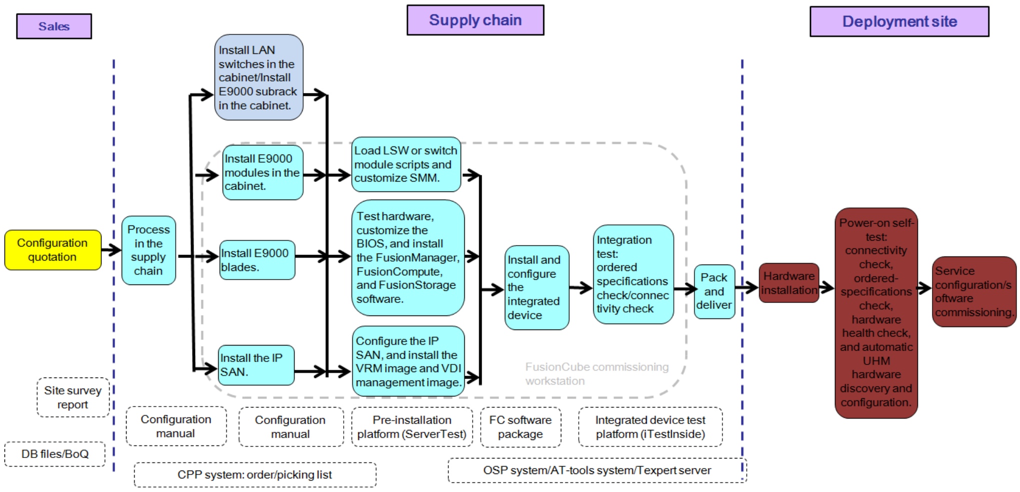
Hardware preinstallation: Devices are installed in cabinets and cables are bundled properly.

Integrated system commissioning: System configuration integrity and connectivity are checked to ensure that the system configuration meets contract requirements and all internal components can communicate with each other.

Preassembled shipment: The cabinet is preassembled and shipped for delivery.

Onsite deployment: After FusionCube is delivered to the site and powered on, users only need to perform a hardware check, completeness check, network interconnection check, and service configuration and commissioning for the system.

Figure 6-3 Integration and delivery process of FusionCube



6.2 Distributed Storage Software

High-performance becomes a key requirement on enterprise storage systems. However, in recent years, continually providing long-term high-performance support for all applications became especially difficult. It has been tested that traditional storage systems can no longer keep up with the demanding storage requirements and increasing data volumes. Moreover, high-end storage devices that serve as independent network elements (NEs) are not only expensive but also need dedicate management and maintenance.

FusionStorage Block, the distributed block storage software running on FusionCube servers, integrates computing and storage and combines local disks on all servers into a virtual storage resource pool. Therefore, it can completely take the place of external SAN devices in some scenarios. FusionStorage Block integrates computing and storage to achieve high performance, reliability, and cost-effectiveness.

Breakthrough architecture and design: The innovative design of FusionStorage Block enables FusionCube to optimize the performance of the traditional architecture. This performance optimization allows users to fully use system resources and enables the system to automatically allocate workloads among all hard disk drives. In addition, FusionStorage Block supports a series of advanced functions, such as thin provisioning and snapshot, without compromising system performance.

Consistent and predictable performance and scalability: The load balancing among all disks and distributed cache structure of FusionStorage Block allow linear and smooth performance improvement by adding server nodes. FusionStorage Block ensures consistent high performance at peak and valley hours regardless of the change of the service amount and snapshot mode or even though the component becomes faulty.

Flexibility and self-healing: FusionStorage Block retains high flexibility and runs properly even if hardware becomes faulty. In addition, the self-healing capability empowers FusionStorage to defend against more hardware faults that may occur after its initial recovery from failures.

Deep convergence of computing and storage: FusionStorage Block integrates the local disks on all servers as a virtual storage resource pool and delivers automatic management, high running performance, and out-of-the-box packaging for quick deployment. FusionStorage Block does not serve as an independent NE for dedicate configuration or management. In this way, resources on the compute nodes can be fully used, which not only provides users with high storage performance, but also reduces costs.

6.2.5 Distributed Engines

FusionStorage Block uses distributed stateless engines. These engines are deployed on each server to prevent performance bottlenecks that may be caused by traditional, centrally deployed engines. Moreover, these distributed engines deployed on standalone servers consume much fewer CPU resources but provide much higher IOPS than centrally deployed engines do.

6.2.6 Distributed Caches

FusionStorage Block integrates computing and storage resources and evenly distributes caches and bandwidths to all servers.

Each disk on FusionStorage Block servers uses independent I/O bandwidths, preventing a large number of disks from competing for limited bandwidths between computing devices and storage devices in an independent storage system.

FusionStorage Block can use certain server memory as the read cache and SSDs as the write cache. Caches are evenly distributed to all nodes. The total cache size on all servers is far greater than that provided by external storage devices. Even when using the large-capacity and low-cost SATA disks, FusionStorage Block can still provide one to three times higher I/O performance and larger effective capacity.

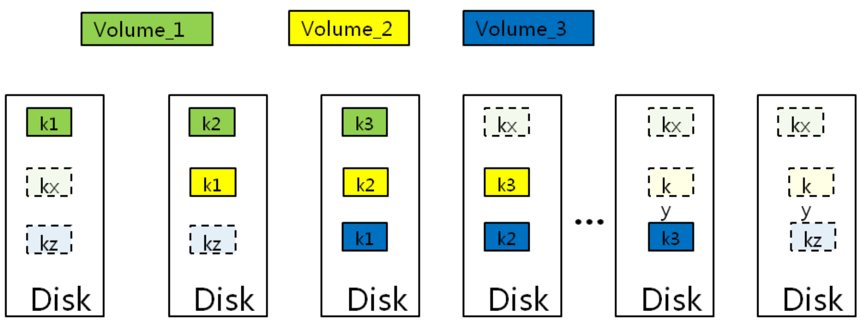
6.2.7 Thin Provisioning

FusionStorage Block provides the thin provisioning function, which allows users to use much more storage space than that actually available on the physical storage device. This feature significantly improves storage space utilization compared with thick provisioning.

FusionStorage Block uses the distributed hash technology to support thining provisioning, without pre-allocating space.

Thin provisioning effectively prevents performance deteriation caused by IP SAN capacity expansion.

Figure 6-4 Thin provisioning of FusionStorage Block



6.3 High-Performance SSD Storage

I/O imbalance becomes noticeable as users and applications increase exponentially as well as the overwhelming development imbalance occurs between CPUs and storage. Data read and write performance must be improved to ensure service stable operating and enhances user experience. The data warehouse applications with massive data analysis impose high requirements for storage throughput.

By communicating with the CPU directly, the NVMe PCIe SSD cards shorten storage access paths and combine flash high-speed read and write performance to significantly improve storage performance and resolve the storage I/O bottleneck. Huawei FusionCube supports distributed SSDs designed for high-performance applications. It provides higher read and write performance than traditional SATA or SAS HDDs.

Huawei SSD cards feature the following:

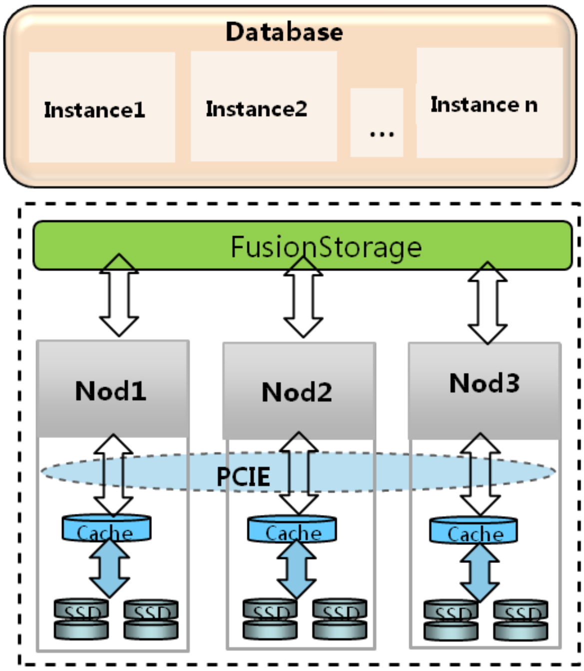
- High bandwidth: uses PCIe 3.0 x4 interfaces and provides up to 3.2 Gbit/s read/write bandwidth.
- High I/O: supports 4 KB data blocks and 100% random I/O, and provides 800,000 continuous random read IOPS and 175,000 continuous write IOPS.

The distributed SSD storage system takes the following measures to enhance system reliability:

- Embedded ECC engine and RAID engine, and a 2D checking/correction mechanism among channels.
- Prevents data errors with an embedded data scrubbing engine that verifies storage data in real time.
- Shares resources between multiple channels using the dynamic RAID algorithm, and ensures normal operations even with multiple faulty chips or bad blocks in a chip.

It prolongs the SSD service life by classifying and managing internal cold and hot data and by implementing advanced wear algorithms that maximize recovery efficiency and reduce write wear.

Figure 6-5 SSD storage system



6.4 High-Speed InfiniBand Switching

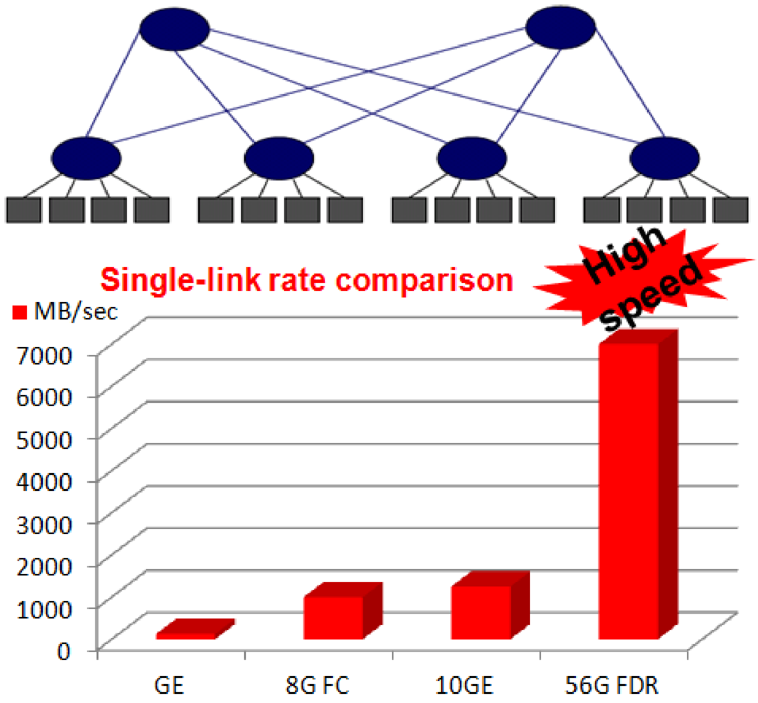
FusionCube uses the industry-leading InfiniBand (IB) network that is based on the edge-cutting remote direct memory access (RDMA) technology, which enhances the database throughput, speeds up response to requests, and improves user experience.

FusionCube uses IB networking for the communication between storage nodes and database nodes (compute nodes), thereby implementing dual-link redundancy and ensuring system reliability.

The FusionCube delivers the following characteristics using the high-speed InfiniBand network:

- Provides a data rate of 56 Gbit/s to allow high-speed interconnection.
- Uses standard multilayer fat-tree networking for smooth capacity expansion.
- Provides a communication network where congestion hardly occurs and avoids data switching bottlenecks.
- Transmits computing and storage information with a delay in microseconds.
- Provides lossless network QoS and ensures data integrity during transmission.
- Allows multi-path communication for active and standby ports and ensures communication path redundancy.

Figure 6-6 High-Speed InfiniBand network



6.5 Network Reliability Design

The FusionCube network is logically divided into the service plane, storage plane, and management plane, which are isolated from each other to protect the infrastructure platform.

Storage plane: Storage devices on the server communicate with each other over layer 2 storage network. Storage devices provide storage resources for virtual machines (VMs) through the virtualization platform but do not communicate with VMs directly.

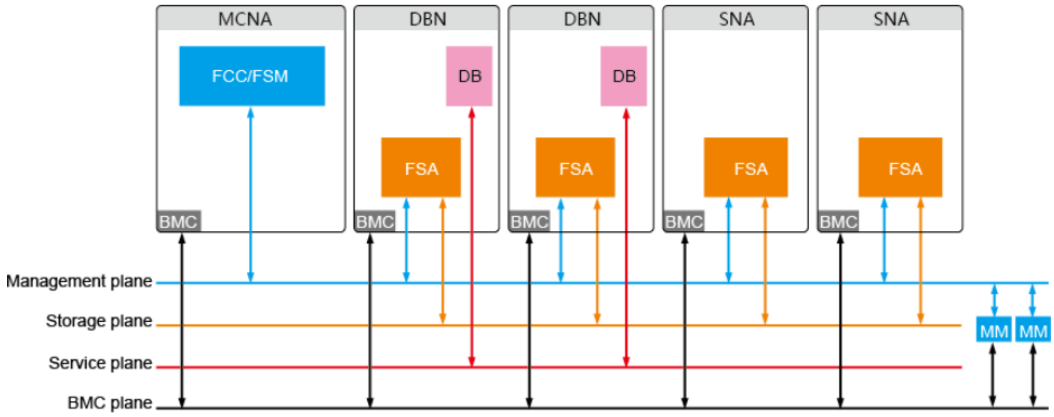
Service plane: provides a channel for users to obtain services, for virtual NICs of VMs to communicate with each other, and for external applications to interact with the FusionCube system. Access can be separated by the VLANs configured for VMs.

Management plane: supports management, service deployment, and system loading of the entire cloud computing system. The BMC plane manages servers and is isolated from the management plane.

Each blade server provides service ports and management ports to ensure link redundancy by NIC binding.

Two 56 Gbit/s IB network ports are deployed for the storage plane to implement dual-link redundancy, thereby enhancing system reliability.

The following figure illustrates the networking pattern.



6.6 Hardware Platform

6.6.7 Blade Servers

6.6.7.1 E9000 Chassis

Huawei E9000 is a 12U blade server that houses a wide variety of compute nodes, including service processing nodes, storage nodes, and resource expansion nodes. The E9000 has the following functions and features:

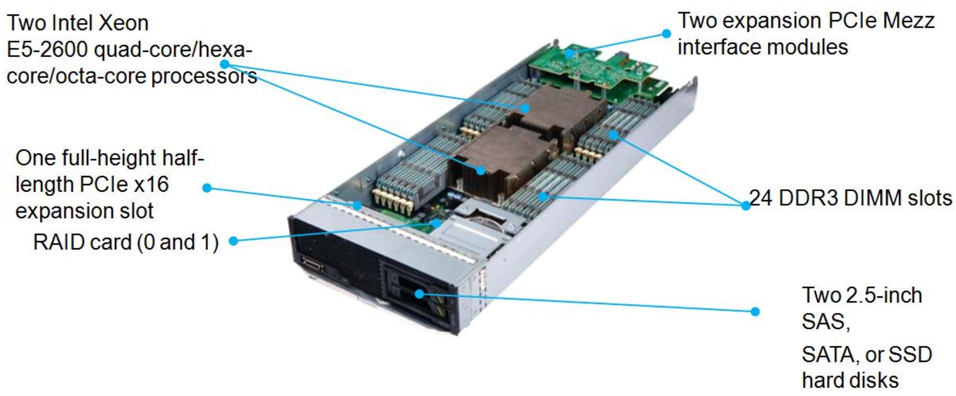
- An E9000 chassis can be configured with up to eight full-width or 16 half-width compute nodes in the front.
- A half-width slot provides cooling capacity of up to 850 W.
- A full-width slot provides cooling capacity of up to 1700 W.
- A half-width compute node supports up to two processors and 24 DIMMs.
- A full-width compute node supports up to four processors and 48 DIMMs.
- A chassis provides the computing density of 32 processors or 256 cores with a maximum memory capacity of 12 TB.
- The midplane supports a maximum switching capacity of 5.76 Tbit/s.
- A chassis provides four (two pairs of) slots for switch modules, which support Ethernet, Fibre Channel (FC), Fibre Channel over Ethernet (FCoE), and IB protocols and provide pass-through I/O ports.

Figure 6-7 E9000 appearance

6.6.7.2 E9000 Blades

FusionCube supports four types of blades: CH121 V3 computing blade, CH242 V3 computing blade, CH222 V3 storage blade, and CH225 V3 storage blade.

Figure 6-8 Appearance of CH121 V3 compute node



The CH121 V3 has the following features:

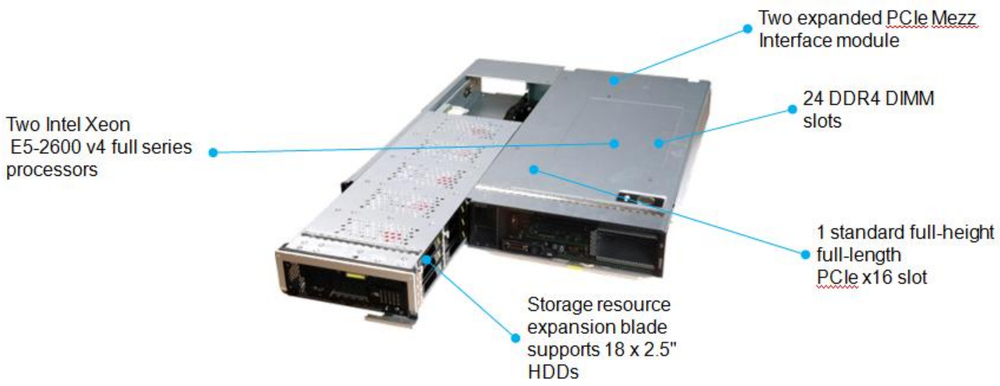
- Supports all Intel® Xeon® E5-2600 v4 series high-performance processors to implement outstanding 44-core computing performance.
- Provides 24 DIMM slots with up to 3 TB DDR4 memory capacity.

Figure 6-9 Appearance of CH242 V3 compute node

The CH242 V3 computing blade has the following features:

2. Uses four Intel® Xeon® E7 v4 processors.
3. Provides 32 DIMM slots with up to 4 TB DDR4 memory capacity.
4. Supports eight 2.5-inch SSD, SAS, or SATA HDDs.
5. Supports two FHHL PCIe x8 standard cards.

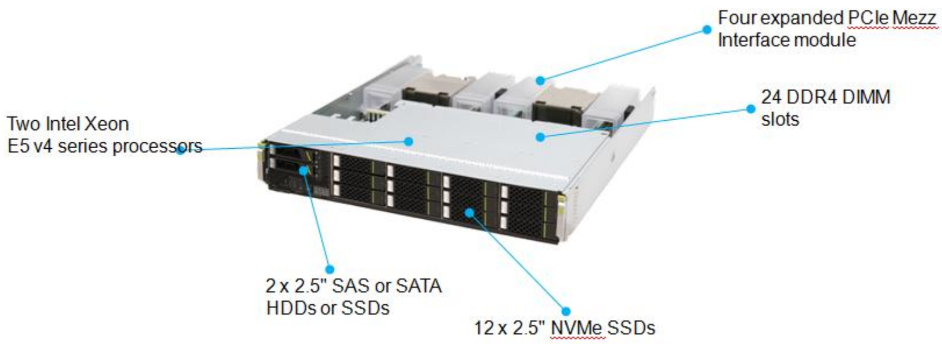
Figure 6-10 Appearance of CH222 V3 storage node



The CH222 V3 has the following features:

6. Uses two Intel® Xeon® E5-2600 v4 processors.
7. Provides 24 DIMM slots with up to 1.5 TB DDR4 memory capacity.
8. Supports up to 15 x 2.5" SSD or SAS/SATA HDDs.
9. Supports hot swap and in-service replacement of a single hard disk. It provides one full-height 3/4-length standard PCIe x16 slot.

Figure 6-11 Appearance of CH225 V3 storage node



The CH225 V3 computing blade has the following features:

10. Uses two Intel® Xeon® E5-2600 processors.
11. Provides 24 DIMM with up to 1 TB DDR4 memory capacity.
12. Supports 2 x 2.5" SSD, SAS or SATA HDDs, and 2 x PCIe SSDs
13. Four mezzanine cards (x16)
14. 12 x 2.5" NVMe PCIe SSDs

6.6.7.3 10GE Switch

FusionCube uses the CX310 switch to support 10GE networks for the service plane and the management plane. Each subrack is configured with two switches. Table 6-1 shows the switching indicators of the CX310.

Table 6-1 Function description of CX310

Item	Description
Appearance	
Number of ports on the panel	Provides one BMC serial port and one SYS serial port. Provides sixteen 10GE uplink optical ports.
Switching capacity	1.28 Tbit/s
Packet forwarding rate	960 mpps
Data center features	Supports TRILL. Supports FCoE. Supports Data center bridging exchanging (DCBX), priority-based flow control (PFC), and enhanced transmission selection (ETS).
VLAN	Supports Access, Trunk, and Hybrid modes. Supports default VLAN. Supports QinQ and enhanced QinQ. Supports MUX VLAN.
QoS	Supports traffic classification policies based on layer 2 protocols, layer 3 protocols, layer 4 protocols, and 802.1p priority levels. Supports access control list (ACL), Control Assessment Review (CAR), Remark, and Scheduling operations. Supports queue scheduling modes, such as priority queuing (PQ), weighted round robin (WRR), dynamic rate repartitioning (DRR), PQ+WRR, and PQ+DRR. Supports congestion avoidance mechanisms such as weighted random early detection (WRED) and tail dropping. Supports traffic shaping.

Item	Description
Security and management	Supports CLI level-based protection to prevent unauthorized access. Defenses against DoS, Address Resolution Protocol (ARP), and Internet Control Message Protocol (ICMP) attacks. Port isolation, port security, and sticky MAC Binding of IP addresses, MAC addresses, port numbers, and VLANs Supports multiple authentication methods, including Authentication, Authorization, and Accounting (AAA), Remote Authentication Dial In User Service (RADIUS), and Huawei Terminal Access Controller Access Control System (HWTACACS). Remote network monitoring (RMON).

6.6.1.4 InfiniBand Switch

FusionCube uses the CX610 IB switch to support 56 Gbit/s data switching for the data plane. Each subrack is configured with two switches. Table 6-2 shows the switch performance indicators of the CX610.

Table 6-2 Function description of CX610

Item	Description
Appearance	
Number of ports on the panel	Provides one BMC serial port and one SYS serial port 18 QDR/FDR InfiniBand QSPF+ ports for uplink
Data center features	Supports TRILL.

6.6.2 Rack Servers

6.6.2.4 RH2288H V3 Server

The Huawei RH2288H V3 server (RH2288H V3 for short) is a 2U two-socket rack server launched by Huawei to meet customer requirements for the Internet, Internet data center (IDC), cloud computing, enterprise market applications, and telecom service applications.

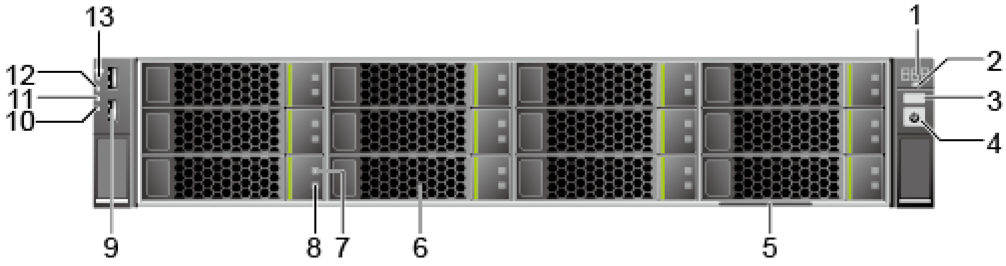
The RH2288H V3 features high-performance computing (HPC), large storage capacity, low power consumption, high scalability and reliability, and is easy to deploy and manage. It is ideal for virtualization, basic enterprise applications, telecom services, and storage services such as distributed storage, data mining, electronic albums, and videos.

Two rear 2.5-inch SAS hard disks. One RAID controller card is required. An RH2288H V3 without NVMe PCIe SSDs supports a maximum of 12 front 3.5-inch SAS HDDs, SATA HDDs, or SSDs.

Front Panel

Figure 6-12 shows the front panel of an RH2288H V3 with 12 hard disks.

Figure 6-12 Front panel of an RH2288H V3 with twelve 3.5-inch hard disks

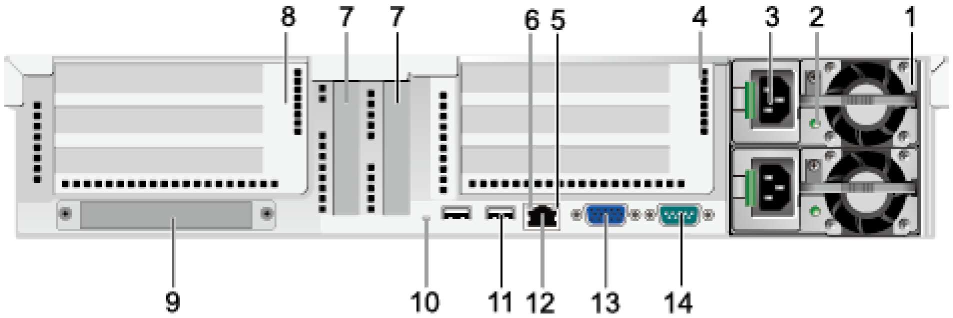


1	Fault diagnosis LED	2	Health indicator
3	UID button/indicator	4	Power button/indicator
5	Slide-out information label	6	Hard disks (numbered 0 to 11 from top down and from left to right)
7	Hard disk fault indicator	8	Hard disk activity indicator
9	USB 2.0 port	10	Network port link indicator 4
11	Network port link indicator 3	12	Network port link indicator 2
13	Network port link indicator 1		

Rear Panel

Figure 6-13 shows the rear panel of an RH2288H V3.

Figure 6-13 Rear panel of an RH2288H V3

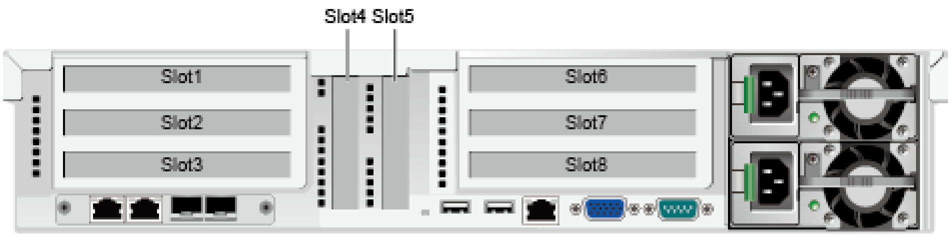


1	PSU	2	PSU indicator
3	PSU socket	4	I/O module 2 or NVMe PCIe SSD transfer module
5	Connection status indicator	6	Data transmission status indicator
7	Onboard PCIe slot	8	I/O module 1
9	NIC	10	UID indicator
11	USB 3.0 port	12	Management network port (Mgmt)
13	VGA port	14	Serial port

PCIe Slot Layout

Figure 6-14 shows the PCIe slot layout of the RH2288H V3.

Figure 6-14 PCIe slot layout



6.6.2.5 RH5885H V3 Server

As technology and applications develop rapidly, customers have higher requirements on the reliability, performance, maintainability, and cost of servers. Building on Huawei's rich experience in servers, the RH5885H V3 is a high-performance enterprise-level rack server that uses the latest Intel processors.

The RH5885H V3 provides higher reliability, flexibility, scalability, and performance than previous servers. To address applications such as databases, virtualization, and memory computing, the RH5885H V3 provides various processing capabilities, memory capacity, and I/O capabilities.

Appearance

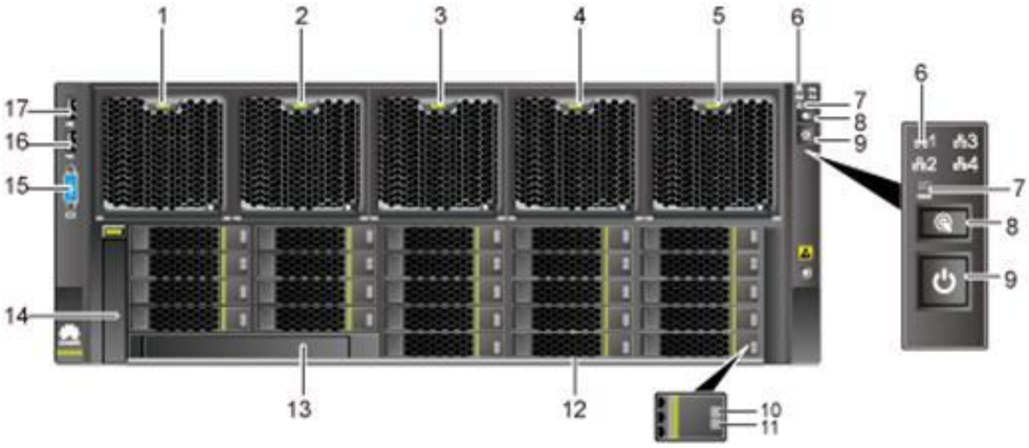
Figure 6-15 shows the appearance of an RH5885H V3.

Figure 6-15 RH5885 V3 appearance

Front Panel

Figure 6-16 shows the front panel of an RH5885H V3 after the front bezel is removed.

Figure 6-16 RH5885H V3 front panel



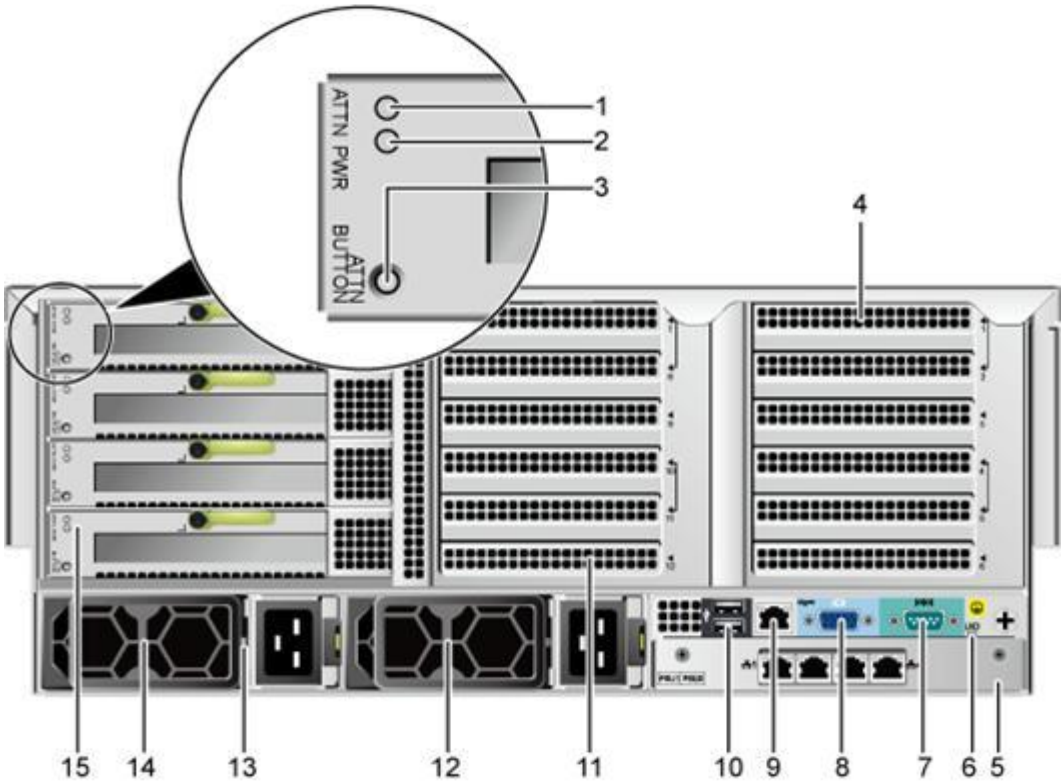
1	Fan module 1	2	Fan module 2
3	Fan module 3	4	Fan module 4
5	Fan module 5	6	Network port link indicator
7	Health indicator	8	UID button/indicator
9	Power button/indicator	10	Hard disk fault indicator
11	Hard disk activity indicator	12	Hard disk
13	DVD drive	14	LCD
15	VGA port	16	USB port 2
17	USB port 1	–	–

NOTE
Unless otherwise specified, this document assumes that the 23 hard disks used by the RH5885H V3 are all 2.5-inch SAS HDDs, SATA HDDs, or SSDs.

Rear Panel

Figure 6-17 shows the rear panel of the RH5885H V3.

Figure 6-17 RH5885H V3 rear panel (two PSUs)

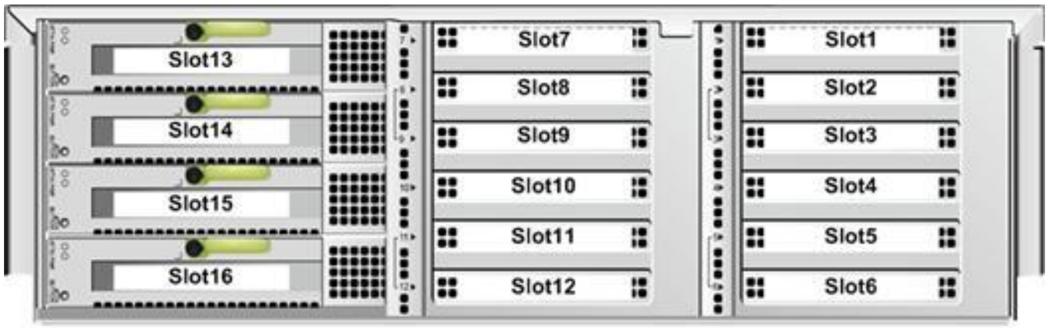


1	Hot swap indicator (hot-swappable PCIe card)	2	Power indicator (hot-swappable PCIe card)
3	Hot swap button (hot-swappable PCIe card)	4	Standard PCIe riser card 1
5	LOM NIC	6	UID indicator
7	Serial port	8	VGA port
9	BMC port	10	USB port
11	Standard PCIe riser card 2	12	PSU 2
13	PSU indicator	14	PSU 1
15	Hot-swappable PCIe riser card	-	-

> **NOTE**
> The RH5885H V3 supports two PSU configuration modes. One is the 1+1 redundancy mode shown in the preceding figure; the other is 2+2 redundancy mode. The latter mode is not shown in the figure. For details about the latter mode, contact Huawei local sales personnel.

Figure 6-18 shows the PCIe slot layout of the RH5885H V3.

Figure 6-18 PCIe slot layout



6.6.2.6 1288H V5 Server

The Huawei FusionServer 1288H V5 (1288H V5 for short) is a vastly versatile 1U 2-socket rack server developed for Internet, Internet data center (IDC), cloud computing, enterprise, and telecom service applications.

The 1288H V5 is ideal for IT core services, cloud computing virtualization, high-performance computing, distributed storage, big data processing, enterprise or telecom service applications, and other complex workloads. It features low power consumption, high scalability and reliability, and easy deployment and management.

Figure 6-19 shows the appearance of a 1288H V5 server with eight hard disks.

Figure 6-19 1228H V5 appearance

1288H V5 has the following features in terms of scalability and performance:

The Intel® Xeon® Scalable Processor ensures high processing performance by providing up to 28 cores, 3.6 GHz clock frequency, 38.5 MB L3 cache, and two 10.4 GT/s UPI links between CPUs.

- Each 1288H V5 supports two processors with 56 cores and 112 threads to maximize the concurrent execution of multithreaded applications.
- L2 cache is added. Each core can exclusively use a maximum of 1 MB L2 cache or 1.375 MB L3 cache.
- Intel® Turbo Boost Technology 2.0 enables processor cores to run at maximum speeds during peak hours by temporarily going beyond the processor thermal design power (TDP).
- Intel Hyper-Threading Technology enables each CPU core to run up to two threads, improving parallel computation capability.
- Intel® Virtualization Technology integrates hardware-level virtualization functions to allow OS vendors to better use hardware to address virtualization workloads.

Twenty-four DDR4 error checking and correcting (ECC) RDIMMs or load-reduced DIMMs (LRDIMMs) provide a maximum memory speed of 2666 MT/s and a maximum memory capacity of 1536 GB, featuring high speed and availability. The maximum memory bandwidth is 249.9375 GB/s in theory.

The 1288H V5 supports flexible hard disk configurations and provides elastic and scalable memory capacities to satisfy storage capacity and upgrade requirements.

The I/O performance of all-SSD configuration is higher than that of mixed configuration of SSDs and HDDs and 100 times that of all-HDDs.

Intel® Advanced Vector Extensions -512 (AVX 2.0) improves floating-point computing performance for computing-intensive applications.

Supports various LAN on motherboard (LOM) cards and flexible network interface cards (NICs) to provide rich network ports.

Supports up to three PCIe 3.0 slots.

The Intel® Xeon® Scalable processors incorporate the PCIe 3.0 controller using Intel Integrated I/O. This remarkably shortens I/O latency and enhances overall system performance.

6.6.2.7 2288H V5 Server

The Huawei FusionServer 2288H V5 (2288H V5 for short) is a 2U 2-socket rack server developed for Internet, Internet data center (IDC), cloud computing, enterprise, and telecom service applications.

Marked H22H-05 on the nameplate, the 2288H V5 is ideal for IT core services, cloud computing virtualization, high-performance computing, distributed storage, big data processing, enterprise or telecom service applications, and other complex workloads. It combines low power consumption with high scalability and reliability, and easy deployment and management.

Figure 6-20 shows the appearance of a 2288H V5 server with 12 hard disks.

Figure 6-19 2288H V5 appearance

2288H V5 has the following features in terms of scalability and performance:

Intel® Xeon® Scalable Processors ensure high processing performance by providing up to 28 cores, 3.6 GHz dominant frequency, 38.5 MB L3 cache, and two 10.4 GT/s UPI links between CPUs.

- Each 2288H V5 supports two processors with 56 cores and 112 threads to maximize the concurrent execution of multithreaded applications.
- L2 cache is added. Each core can exclusively use a maximum of 1 MB L2 cache or 1.375 MB L3 cache.
- Intel® Turbo Boost Technology 2.0 allows CPU cores to run faster than the Thermal Design Power (TDP) configuration specified frequency if the CPU cores are operating below power, current, and temperature specification limits.
- Intel Hyper-Threading Technology enables each CPU core to run up to two threads, improving parallel computation capability.
- Intel® Virtualization Technology integrates hardware-level virtualization functions to allow OS vendors to better use hardware to address virtualization workloads.

Twenty-four DDR4 error checking and correcting (ECC) RDIMMs or load-reduced DIMMs (LRDIMMs) provide a maximum memory speed of 2666 MT/s and a maximum memory capacity of 1536 GB, featuring high speed and availability. The maximum memory bandwidth is 249.9375 GB/s in theory.

The 2288H V5 supports flexible hard disk configurations and provides elastic and scalable memory capacities to satisfy storage capacity and upgrade requirements.

Intel® Advanced Vector Extensions -512 (AVX-512) improves floating-point computing performance for computing-intensive applications.

The I/O performance of all-SSDs is higher than that of mixed configuration of SSDs and HDDs and 100 times that of pure HDDs.

Supports LOM and flexible NIC to provide rich network ports.

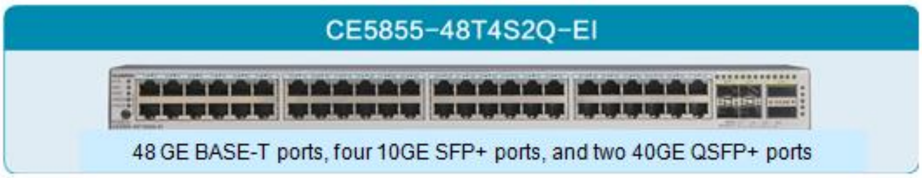
Supports up to eight PCIe 3.0 slots.

The Intel® Xeon® Scalable processors incorporate the PCIe 3.0 controller using Intel Integrated I/O. This remarkably shortens I/O latency and enhances overall system performance.

6.6.2.8 GE Switch

FusionCube uses the CE5885 switch to support GE networks for the service plane and the management plane. Each subrack is configured with two switches. Figure 6-21 shows the switching performance indicators of the CE5855.

Figure 6-21 CE5855-48T4S2Q-EI



6.6.2.9 InfiniBand Switch

FusionCube uses the external Mellanox InfiniBand switch to support 56 Gbit/s high-speed data switching for the data plane. Figure 6-22 shows the switch performance indicators of the Mellanox InfiniBand.

Figure 6-22 MELLANOX-MSX6036F-2SFS